UNIVERSITY OF CALIFORNIA

Santa Barbara

Lengthening Life

Diapause Induced Lifespan Extension in Drosophila melanogaster

A Thesis submitted in partial satisfaction
of the requirements for the degree
Master of Arts in
Molecular, Cellular & Developmental Biology

By

Dominique Houston

Committee in Charge
Professor Denise Montell, Chair
Professor Julie Simpson
Professor Kathleen Foltz

March 2019

The thesis of Dominique Houston is approved.

_____

Kathleen Foltz

_____

Julie Simpson

_____

Denise Montell, Committee Chair

January 2019

Table of Contents

Acknowledgments

"Life seems but a quick succession of busy nothings."
– Jane Austen

"Dream as if you'll live forever, live as if you'll die today."
– James Dean

"What would life be worth if there were no death? Who would enjoy the sun if it never rained? Who would yearn for the day if there were no night?"
– Glenn Ringtved

"We all die. The goal isn't to live forever, the goal is to create something that will."
– Chuck Palahniuk

If I could only take one thing from my experience in the Masters program at UCSB, it is this; to always make friends feel like family wherever you go. Without the influence, advice, and support, from my peers and colleagues, I do not think I could've survived my graduate education. Here's to all of you friends! With a very special and heartfelt mention of my best lab buddy Christopher Conner. You are the most honest and sincere person I've ever had the fortune to meet. You made work fun despite all the failures, and still managed to work harder and smarter than most people I know. We both suffered ups and downs together (for science!) and I greatly appreciated all of our therapy sessions, at least you know you'll always have that career as a fall back.

Another group of amazing and accomplished individuals, that deserve particular mention, are the members of my committee. I am so grateful and proud to have 3 driven, intelligent and kind women agree to suffer through my attempt at a thesis paper. Not only do I acknowledge and thank you for contributing to my academic career, but also in taking a more personal interest in myself. There was never a time when I felt I couldn't talk to any one of you about just about anything, and I am forever grateful for your contributions that helped me make my decision of whether to continue with my graduate education or not. Just being around such passionate people has been both inspiring and motivating. Here's to you ladies for constantly kicking butt!

Finally, I wish to acknowledge my closest companion, Cat. You are the best friend a person could ever ask for. Somehow you always know when I need your cuddles the most and give them generously. If I could make any creature live longer it would be you, I fear the day when you will be gone from me. So, thank you for being my support animal and so much more!

>^_^<

Abstract

Lengthening Life: Diapause Induced Lifespan Extension in *Drosophila melanogaster*

By

Dominique Michele Houston

Diapause is a protective trait used by many invertebrates to overcome unfavorable environmental conditions, such as cold temperature and starvation. *Drosophila melanogaster* undergoes a shallow reproductive diapause, characterized by metabolic reprogramming, reduced locomotion, increased stress tolerance, and delayed tissue senescence, ultimately, resulting in an increase in overall organism lifespan. Several pathways implicated in longevity also regulate diapause. This, along with the prevention of normal degeneration seen during this process, offers a unique way in which to study aging and its associated degenerative diseases. Most studies of this complex trait have utilized the pause of ovary development at pre-vitellogenic stages as the main determinant for categorizing diapause inducibility. This approach ignores the complexity of this trait by only relying on reproductive development as the defining characteristic. In this study, the power of the Drosophila Genetic Reference Panel is implemented in order to conduct a genome-wide association study of diapause traits. By quantifying two phenotypes, fecundity and longevity, this tool has led to the identification of a total of 151 genes associated with this trait, and 52 intergenic regions, which may be involved in the regulation of the diapause program. Enrichment analyses suggest the involvement of metabolic pathways, the innate immune system, and GPCR signaling, components of which have also been shown in *C. elegans*. Further studies should aim to better combine the two mechanisms of quantifying this phenotype, and functionally confirm the involvement of the candidate genes.

Chapter I: Background

Longer Living: Aging and Degeneration

One definition of aging is the degeneration of tissues over time resulting in physical disabilities, cognitive loss and ultimately, death. As the quality of life in first world countries continues to improve, so does the average life expectancy. In 2015, fourteen percent of the global population was age eighty years or older, and by 2055 the population of the older generation is projected to increase from fourteen to twenty percent (Figure 1) (Anon 2015). Naturally, with this increase in longevity comes an increase in the prevalence of age related degenerative diseases as we begin to outlive our vessels. The World Health Organization (WHO) estimated in 2016 that forty-one million deaths were the result of noncommunicable diseases (NCDs), the majority of which were due to the top three NCDs: cardiovascular disease (17.9 million deaths; 44% of NCDs), cancer (9.0 million deaths; 22% of NCDs), and

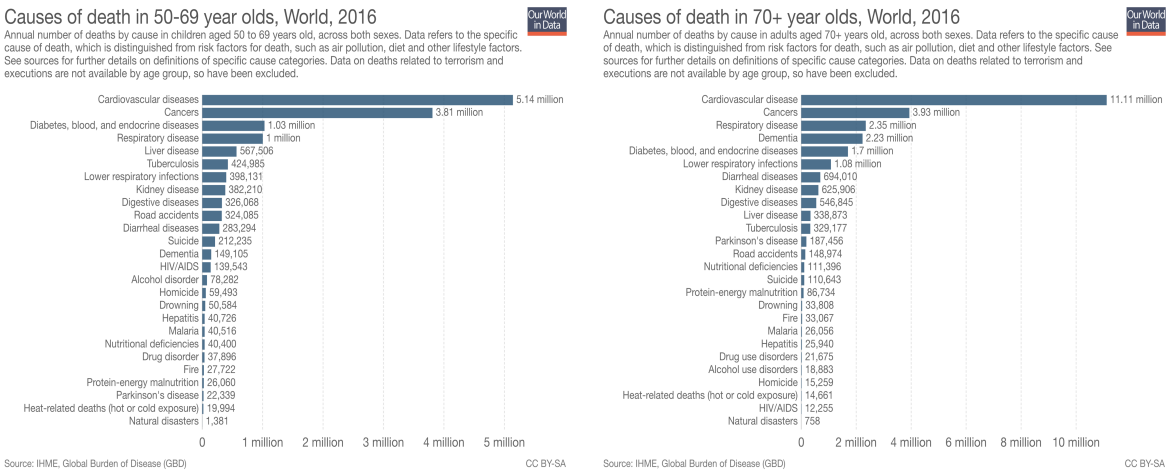

Figure 2 Causes of Death in the Older/Oldest Populations
The top 4 causes of death in the Older and Oldest Populations of people are all Non-communicable diseases attributed to aging. With an increase in this age group globally, it is only natural to see an increase in the incidence of deaths by NCDs associated with aging. Figure adapted from Lee and Mason 2017.

chronic respiratory disease (3.8 million deaths, 9% of NCDs) (World Health Organization 2018; Roser and Ritchie 2018). To date, non-communicable diseases are the leading cause of

death in persons aged sixty and above (Figure 2), and are predicted to cost the global economy $47 trillion in prevention and care by 2030 (Lee and Mason 2017) . Thus, the study of aging and age-related diseases (gerontology) has become a hot-topic in scientific research, and a favorite for funding organizations. This has lead to an increasing demand for the discovery/development of novel methods for studying the many encompassed sub-phenotypes of aging, with the goal of uncovering the molecular mechanisms involved.

Theories & Hallmarks of Aging

To address this issue, we must first consider what is currently known about the aging process. The initial theories describing the progression of age, known as the Fundamental Limitation Theories (wear and tear theories), were based primarily on the Darwinian notion of "survival of the fittest". These theories posited that biological aging is simply the product of universal deteriorative processes. With the generation of new data and insights, these theories have fallen out of favor in the research community. The most recognized theories of aging, fall into two overarching ideas. The first is that aging is a programmed event (evolutionarily conserved), while the second suggests that aging is a result of accumulated cellular and molecular damage (mutation accumulation), though they are not mutually exclusive (Sergiev et al. 2015). A historical timeline of the development of

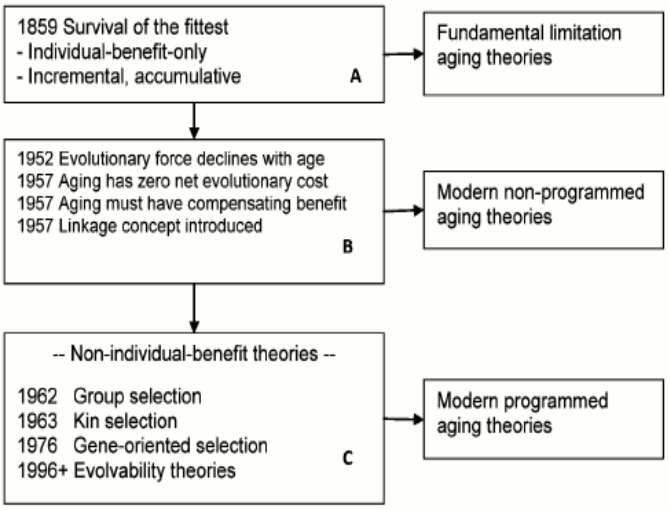

Figure 3 Historical Timeline of Aging Theories This figure depicts the component theories developed overtime that resulted in the initial and present day theories of aging. Adapted from Goldsmith, T.C. (2014).

these theories can be seen in Figure 3. The programmed theories focus on the idea that it is beneficial for post reproductive animals to die to liberate resources for the young. Thus, organisms may have evolved mechanisms that limit lifespan, and in support of this theory, is the fact that there is low variability in lifespan within species.

The key concepts behind the non-programmed theories of aging are that an organism possesses an innate ability to combat typical deteriorative processes, and that any given species has an evolutionary drive to achieve a certain minimum lifespan specific to each species (Goldsmith 2014). Each of these blanket theories contain a multitude of sub-theories more specific to individual hallmarks associated with aging for more information on the sub-theories of aging, please see the following works (Fedarko 2018; Jin 2010). A recent review on "The Hallmarks of Aging" attributes 9 to aging (Figure 4) in various organisms, with an emphasis on work in mammals (López-Otín et al. 2013). These Hallmarks include: Telomere Attrition, Genomic Instability, Epigenetic Alterations, Mitochondrial Dysfunction, Altered Intercellular Communication, Deregulated Nutrient Sensing, Loss of Proteostasis, and lastly, Cellular Senescence. Undoubtedly these hallmarks will likely be updated and appended, as research in the field progresses.

A key question remaining in the field of gerontology is how aging on a cellular and histological scale relates to the overall aging of the organism. Dissecting this process into subsets of replicative aging, cellular aging, tissue aging and organismal aging, may help to answer this question. By distinguishing which hallmarks are more likely the initial contributors, from those that are resultant, consequential contributors, perhaps we can better dissect the complex web of integrated signals and events that comprise the aging process. A small selection of the hallmarks and their relevant findings are discussed below. For an extensive review on the involvement of epigenetic alterations and aging, please see Pal and Tyler (Pal and Tyler 2016).

Figure 4 The Nine Proposed Hallmarks of Aging
Each Hallmark represents an area where homeostasis, or normal regulation are malfunctioning, resulting in biophysiological deterioration. Adapted from Lopez-Otin, C. et al (2013).

Mitochondrial Health & ROS

Mitochondria, the power house of the cell. The main roles of these bacterial descendants are (i) to produce energy for the cell in the form of ATP via the citric acid or Kreb's cycle, (ii) regulate cellular metabolism, and (iii) modulate apoptosis. During normal cellular respiration and metabolism, reactive oxygen species and other metabolic byproducts are generated (Li et al. 2013). Therefore, most theories of the mitochondrial contributions in aging revolve around the generation of ROS during aerobic respiration and the decline of

mitochondrial health due to DNA (mtDNA) damage. While severe dysfunction of mitochondria has been shown to lead to premature aging and lethality, several experiments in multiple model organisms demonstrated that mild inhibition of mitochondrial respiration promotes longevity (Kenyon 2010). RNAi screens of the Electron Transport Chain (ETC), see Figure 5, in *Drosophila* and *C. elegans* have identified multiple genes that impair mitochondria function, but have noted that this effect is independent of lower ROS (Lee et al. 2003; Copeland et al. 2009). Conservation of this mechanism has also been shown between worms and mammals using mutants of an enzyme essential in ubiquinone (coenzyme Q) biosynthesis, a key component in the ETC (Braeckman et al. 1999; Liu et al. 2005). For an in depth review on the contribution of mitochondria to longevity see (Hwang et al. 2012).

Strong evidence supports the idea that ROS are detrimental to cells and the overall life of the organism. Experiments in flies and mammals have suggested that exposure to increased ROS inactivating enzymes results in lifespan extension, though countering evidence exists, suggesting that these effects on longevity may be species dependent (Sergiev et al. 2015). For example, an experiment in mammals has demonstrated a negative correlation between the level of ROS metabolizing enzymes and longevity (Barja 2002). However, this could be due to the presence of these ROS metabolizing enzymes operating in a feedback loop, causing the activation of repair or protective pathways that in turn are detrimental to the cell, or the enzymes themselves could be directly detrimental in the absence of ROS. Regardless, between species there are many conflicting results and so it must be considered that a species-specific influence is present, and perhaps even cell type specificity as well. To further complicate matters, the type of ROS, as well as the subcellular location of antioxidant activity, as opposed to merely the level of ROS, can strongly affect

whether increased ROS correlate with decreased lifespan (Mookerjee et al. 2010). It is likely that a correlation exists between the duration of increased stress response, such as that for ROS, and the resulting effect on longevity. This would suggest that short term exposure to environmental stresses can be beneficial, depending on the type of stressor, while longer exposures result in detrimental effects (Epel and Lithgow 2014; Zhou et al. 2011).

Insulin/IGF & Target of Rapamycin Signaling

In the 1980s, the classical Insulin/IGF-1 (insulin like growth factor) pathway was found to regulate the aging process through the discovery of *age-1*(PI3K subunit) *Caenorhabditis elegans* (*C.* elegans) mutants (Klass 1983). Elucidation of many other components in the Insulin/IGF-1 (IIS) pathway soon followed, such as *daf-2* the *C. elegans* insulin receptor homolog (Kenyon 2011). Since the initial isolation of those first long-lived

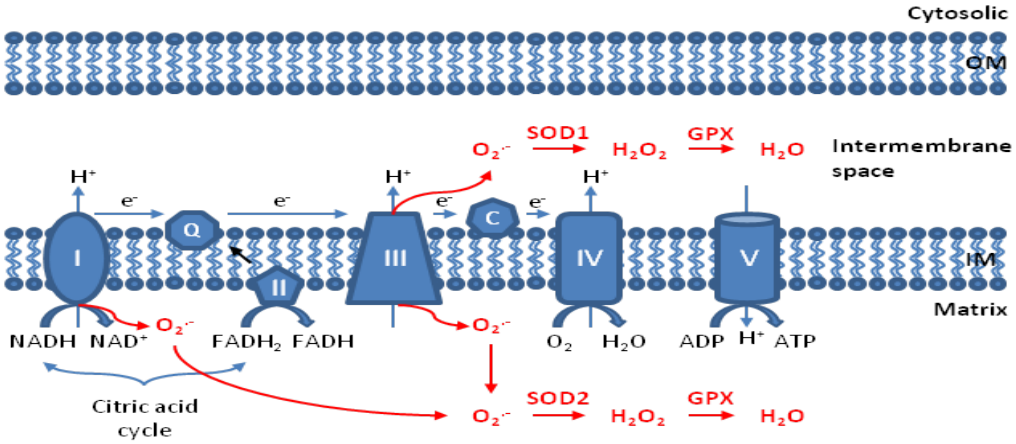

Figure 5 The ETC and Mitochondrial ROS
Metabolites are oxidized by NADH and $FADH_2$, donated electrons are passed down the ETC and ultimately reducing O_2 to produce H_2O. ROS result from electron leakage which form superoxide (O_2^-). This free radical can then be dismutated to hydrogen peroxide (H_2O_2) by superoxide dismutase (SOD1/2) on either side of the inner membrane. The hydrogen peroxide can be further reduced by glutathione peroxidase (GTX), neutralizing the reactive species. Adapted from Li, X. et al (2013).

mutants, 30 years of subsequent research has shown in model organisms from yeast to mammals, that inhibition of IIS signaling (caloric restriction), at multiple levels, results in increased longevity (Fontana et al. 2010; Houtkooper et al. 2010; Newgard and Pessin 2014; Broughton and Partridge 2009). When circulating insulin protein binds to the Insulin Receptor (InR), a receptor tyrosine-kinase, several tyrosine residues are autophosphorylated, promoting interaction with multiple adaptor proteins, such as members of the Insulin Receptor Substrate (IRS) family. Two main pathways emanate from the activated InR node, the Phosphoinositol-3-kinase (PI3K) pathway, and the Extracellular Signal Regulated Kinase (ERK) pathway (Figure 6). The former is responsible for many of the metabolic effects of IIS and only acts through the interaction with IRS, while the latter regulates gene expression, and controls cell growth and differentiation via IRS and Shc adaptor protein transduction cascades. In the PIP3K pathway, activated IRS proteins are then recognized by PI3K, which converts phosphatidylinositol (4,5) bisphosphate (PIP2) into phosphatidylinositol (3,4,5) triphosphate (PIP3). PIP3 recruits Akt, a serine/threonine kinase, to the membrane where it is phosphorylated first by mTORC2 and then by PIP3 dependent Protein Kinase-1 (PDK1). Active Akt then acts in the cytoplasm to activate and inhibit several downstream targets, ultimately resulting in the promotion of lipid, protein and glycogen synthesis, as well as cell growth and gene expression (De Meyts 2000). One critical downstream target of Akt is the Forkhead Box Transcription factor FOXO. When irreversibly phosphorylated by Akt FOXO is inactive in the cytoplasm and eventually degraded. When IIS signaling is attenuated, FOXO remains active and translocates into the nucleus and promotes secondary metabolism such as gluconeogenesis and depresses cell cycle progression (De Meyts 2000). This pinnacle transcription factor has now been shown to regulate several genetic programs upon

diverse post-translational modifications allowing the interaction with numerous FOXO binding partners. Several independent studies on FOXO variants in humans found that a single-nucleotide polymorphism (SNP) in one of the 4 mammalian FOXOs (FOXO3A) associates with increased longevity (Martins et al. 2016). While a multitude of overexpression and repression studies have solidified its lifespan extending role in worms and flies (Davy et al. 2018; Yamamoto and Tatar 2011; Hwangbo et al. 2004).

Another crucial component affected by Akt signaling, is the target of Rapamycin

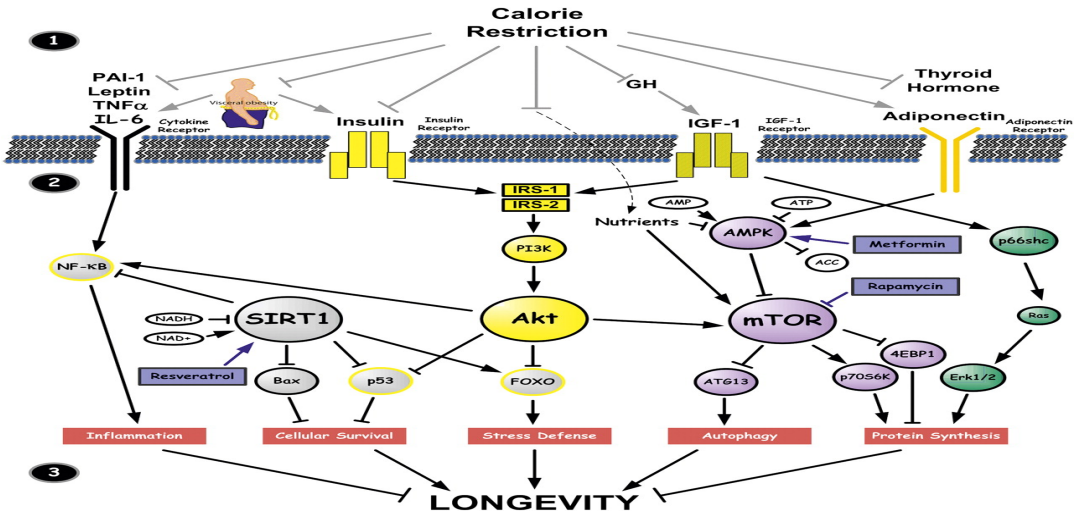

Figure 6 Main Metabolic Pathways of Longevity
Simplified overview of the main metabolic pathways effecting longevity and their integrative nodes. Adapted from Barzilai, N. et al (2012).

(mTOR) pathway (Figure 6). The mTOR pathway, was initially identified in a screen with the drug Rapamycin. Akt directly inhibits the tuberous sclerosis complex (TSC1/2), which is itself an inhibitor of mTORC1, thus allowing mTORC1 dependent signaling to occur. mTORC1 and mTORC2, are stimulated by multiple nutritional and environmental stimuli, such as amino acids, hormones and oxygen. A great deal has been discovered in relation to mTORC1 signaling and its downstream targets, while many mysteries remain for mTORC2 (Newgard and Pessin 2014; López-Otín et al. 2013; López-Otín et al. 2016). Briefly, activation of mTORC1 promotes anabolic activities of the cell such as, mRNA translation,

mitochondrial biogenesis and lipid synthesis, and limits catabolic activities like autophagy (Laplante and Sabatini 2009). Downregulation of mTORC1 activity in yeast, worms and flies, has extended longevity, and increases these lifespan benefits when coupled with dietary restriction (Johnson et al. 2013). This and other evidence indicates that, despite the overarching nature of these two metabolic pathways, there are multiple roads on the journey to "eternity". Given that these nutrient signaling pathways are sensors of the present environmental status, it is of no surprise that they represent the central regulators of aging, and disentangling their complex web of interactions (Figure 6) continues to prove a daunting task. For a concise yet detailed review of the metabolic pathways of aging, see the essay published in 2010 and the updates from 2012 and 2014 (Houtkooper et al. 2010; Newgard and Pessin 2014; Barzilai et al. 2012).

Suspended Animation: Delayed Development in Diapause

Diapause is a protective trait found in many invertebrates (and some vertebrates), and can be found across all stages of development; embryonic, larval, pupal and adult. Not only does this divergence into a delayed developmental state allow an organism to overcome unfavorable environmental conditions (i.e. cold temperature and famine), but also suspends senescence, resulting in an increase in lifespan, in some cases 10-fold, and delayed tissue degeneration (Hand et al. 2016; Tatar et al. 2001; Tatar and Yin 2001). Diapause comes in two main flavors, obligative and facultative. Obligate diapause occurs as a pre-determined developmental program independent of environmental influence, while facultative diapause is triggered by external environmental changes, such as shortened photoperiod, cold

temperature, or starvation (Denlinger 2002; Schiesari and O'Connor 2013). Generally, diapause is described by several stages: sensation; initiation; maintenance; termination; and recovery. Each stage has specific molecular events that have been studied in numerous model organisms such as, *Nematodes, Insects, Crustaceans* and some *Teleosts* (Hand et al. 2016; Salminen et al. 2015; Kankare et al. 2010; Kostál 2006). Studies of diapause across these organisms have implicated signaling pathways known to also play roles in longevity, metabolism, and cell cycle regulation (Sim and Denlinger 2013; Sim and Denlinger 2008; MacRae 2010; Liu et al. 2017; Hand et al. 2016; Tammariello 2001). It is for these reasons that this trait poses a novel means through which to dissect the complex phenotype that is aging, or rather, uncover a way to delay the penultimate degeneration of the body before death.

Diapause in C. elegans

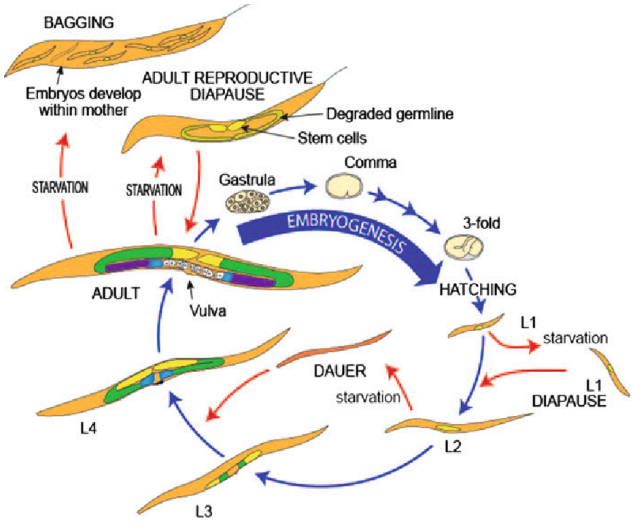

Figure 7 C. elegans Development and Diapause Stages Overview of C. elegans development and larval transitions. C. elegans possess an L1 diapause, the L2 or dauer diapause, as well as an adult reproductive diapause. Figure adapted from Pazdernik and Schedl (2013).

As with longevity, a vast amount of our knowledge of diapause comes from the nematode *Caenorhabditis elegans*. Under conditions such as starvation or overcrowding, *C. elegans* undergoes a distinct form of dormancy, known as dauer, when transitioning from its second larval stage to its third, the L2 to L3 transition (Fielenbach and Antebi 2008; Karp 2018). Several other diapause states exist between additional transitional stages (Figure 7). In particular, interest in the L1 diapause state is increasing due to its relevance to aging,

diabetes, obesity, and cancer (Baugh 2013). During the late L4 stage, if the larva is subject to starvation, it can undergo an adult reproductive diapause (ADR) (Angelo and Van Gilst 2009; Seidel and Kimble 2011; Hand et al. 2016). During this reproductive dormancy, the germ line of the animal undergoes programmed cell death (PCD), and through some unknown mechanism confers protection of the germ line stem cells, maintaining reproductive capacity post-diapause. When worms deficient in the *ced-3* gene (a caspase essential for apoptosis) were subjected to ARD conditions, they were able to initiate ARD, but were unable to reduce the germ cell nuclei normally observed during this process, leading to an inability of these mutants to fully recover and reproduce (Seidel and Kimble 2011; Angelo and Van Gilst 2009). Each of these arrested stages are most easily differentiated by their duration, but can also be distinguished by their unique molecular events. Current research in the field aims to identify the differences and similarities between each of the diapause states to better understand these phenomena.

Decades of diapause research in *C. elegans* has focused on the dauer stage, and has identified several important signaling pathways involved in the process. A detailed review (Fielenbach and Antebi 2008) summarizes the sensory and signaling pathways involved in dauer induction, maintenance and recovery. The sensation of dauer/diapause inducing conditions occurs through several environmental stimuli. These cues can be food prevalence, temperature, photoperiod, and pheromones, and ultimately integrate to direct the organism down the correct developmental path. The most efficacious inducer of dauer formation is the dauer pheromone, which is secreted during periods of overcrowding and low food resources. In 2005 the structure of this pheromone was elucidated and found to be a family of 3 related types of ascarosides, glycolipids containing ascarylose sugar, with the most potent (2) acting

in the nanomolar range (Jeong et al. 2005; Butcher et al. 2009). Despite over 25 years of research into the mechanisms of dauer pheromone action, much remains a mystery. Sensation of these unfavorable environmental cues, integrates at two main sensory organs, called amphids, in the head of the worm. Ablation of specific subsets of cells in these organs (ASI, ASF and ASG), can lead to transient dauer formation regardless of environmental conditions. Furthermore, ablation of ASJ led to continual developmental arrest, indicating the role of these cells in maintaining the normal development program, and promoting recovery from dauer (Fielenbach and Antebi 2008). The role of amphids in dauer pheromone detection has been further demonstrated when strains carrying loss of function mutations in *srbc-64* and *srbc-66* (GPCRs), primarily expressed in ASK, displayed strong defects in dauer formation that varied in response to the individual ascarosides (Kim et al. 2009). However a strain carrying null mutations in both genes was still capable of dauer initiation, suggesting that other receptors remain and, given the >1000 orphan GPCRs, are likely still in this receptor family (Sommer and Ogawa 2011; Fielenbach and Antebi 2008). In addition to their roles in chemosensation, these cells also secrete hormones; perfectly poising them as orchestrators of sensory transmittance, and endocrine signaling, involved in dauer formation.

Historically, loci involved in the dauer process were categorized into pathways based on genetic epistasis and complementation experiments. Those identified involved components of sensory neurons, cyclic guanosine monophosphate (cGMP) signaling, as well as neuronal serotonergic signaling. All of these have been proposed to impinge upon TGF-beta and insulin/IGF-1 signaling, which then collectively converge on steroid hormone receptor transcriptional cascades, crucial in initiating the dauer decision. Seven genetically grouped loci, *daf-1, daf-4, daf-7, daf-8, daf-14, daf-3* and *daf-5,* were all found to influence

dauer formation, and all are components of the TGF-ß pathway. Under normal conditions this pathway is activated and consequent signaling promotes reproductive growth and hormone biosynthesis (Figure 8). In addition, TGF-beta signaling has been shown to regulate insulin like proteins (ILPs), and numerous studies have shown significant crosstalk between these two pathways (Budi et al. 2015; Xue et al. 2012; Guo and Wang 2009; Fielenbach and Antebi 2008; Tatar 2010). For example, TGF-beta *daf-c* mutants are suppressed partially by *daf-16*/*FOXO* mutants, the downstream effector of IIS signaling. Furthermore, the downstream effector of TGF-beta signaling DAF-3/SMAD, shares many putative target genes with DAF-16/FOXO, and expression profiling has shown that they often have associated binding sites (Liu et al. 2004). Additional involvement of IIS has been demonstrated using *Daf-2/InR* mutants that show spontaneous initiation of dauer under normal conditions and, as seen with caloric restriction, have increased longevities (Hand et al. 2016). While much has been discovered with respect to the key signaling pathways involved in dauer formation, it is still unknown how the sensation of dauer inducing stimuli is integrated and orchestrated to result in the dauer decision. Furthermore, the ultimate downstream targets of the signaling pathways involved remain elusive, particularly the

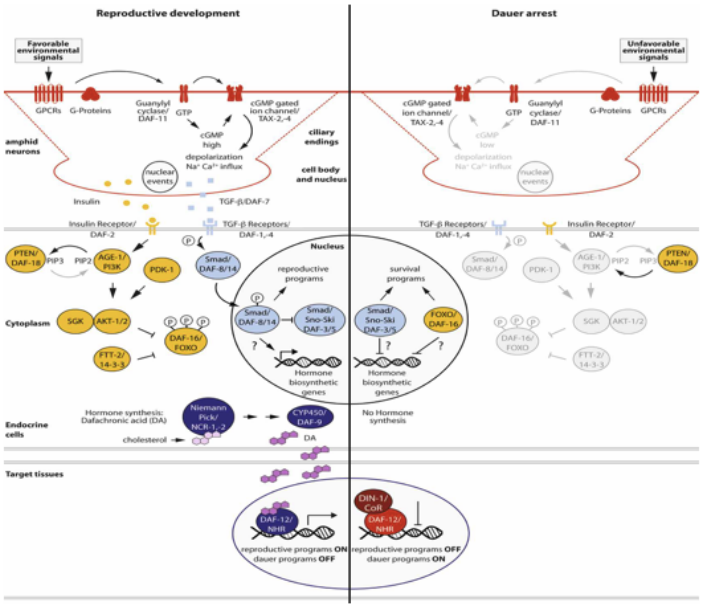

Figure 8 Overview of Proposed Signaling Involved in Dauer Formation During favorable conditions insulin and TGF-ß signaling is active promoting growth, cell divisions and reproduction. When overcrowding or starvation occurs signaling through these pathways is decreased resulting in cell cycle arrest and delayed reproductive development. Figure adapted from Fielenbach and Antebi (2008).

proposed nuclear hormone receptor that likely enacts a huge alteration of genetic architecture and its resulting transcriptome.

Diapause in Drosophila melanogaster

Distinguishing characteristics of diapause induction in *Drosophila* are similar to those seen in other organisms, and involve delayed tissue and organismal senescence, decreased feeding and metabolic activity, increased immune response, decreased locomotion, and delayed reproductive development (Kubrak et al. 2014). Unlike other organisms, there is currently no known pheromone associated with diapause induction in *Drosophila*, and it is rapidly reversed upon removal from diapause conditions (Saunders et al. 1989; Allen 2007; Kubrak et al. 2016; Anduaga et al. 2018; Kubrak et al. 2014). Due to this rapid recovery upon removal from cold temperatures and short-day length (10˚C-15˚C, 8-10L:16-14D), *D. melanogaster* is said to undergo a shallow reproductive diapause. This has led some to the conclusion that Drosophila melanogaster does not undergo a "true diapause", where the organism enters a different developmental state, but rather a "cold quiescence" where normal development is simply halted (Zonato et al. 2017; Saunders et al. 1989). Since the reproductive dormancy of *D. melanogaster* possesses qualities of both diapause and quiescence, it shall hence be referred to as diapause. Several studies have thoroughly examined and characterized this trait in *Drosophila* through physical characterization of tissue changes as well as transcriptional profiling (Salminen et al. 2015; Baker and Russell 2009; Zhao et al. 2016; Kankare et al. 2016; Kučerová et al. 2016). Physical characterization of tissues during diapause has highlighted that systemic effects of diapause induction are great, and likely unique to each individual tissue. Transcriptional studies have highlighted the

changes that occur in carbohydrate and lipid homeostasis, regulation of gene expression related to IIS signaling, and stress response during diapause, like those seen in *C. elegans*.

One of the most interesting components of diapause is the maintenance of individual tissues during this time. Several labs have studied this trait from a physiological standpoint and examined tissues such as the gastrointestinal system, and the neuromuscular system. One group found that during diapause the ISCs (intestinal stem cells) decrease in number, based on fluorescent labeling. Whether this is due to actual loss of intestinal cells or a decrease in expression of the IPC marker remains unclear. This same study also reported physical changes in gut morphology that persisted after recovery from diapause for the remainder of the flies' lifespans, such as a decrease in crop size (Kubrak et al. 2014). Another group observed the neuroprotective effect of diapause by performing behavioral assays of sleep, negative geotaxis and exploratory walking. They showed that senescence of these behaviors was reduced in dormant flies, and that morphology changes of specific neuromuscular junctions seen during aging were not present in the post-diapause cohorts (Liao et al. 2017). These findings suggest that resident stem cells in various tissues undergo preservation, appearing to enter a quiescent state. This protective effect on a systemic level has numerous implications in disease modeling and potential remediation.

Many studies of diapause in *Drosophilids,* and other insects, have utilized transcriptional profiling during diapause to gain insight into its dynamics (Zhao et al. 2016; Ragland et al. 2010; Salminen et al. 2015; Kankare et al. 2010; Kang et al. 2016; MacRae 2010; Kankare et al. 2016). One group found more than 4500 differentially regulated genes upon diapause induction by measuring whole body transcripts of three week diapausing flies and comparing to one week controls. This group found that overall there was a decrease in

IIS, TOR signaling, and MAPK signaling, and an increase in Toll dependent immune signaling, JNK signaling and JAK/STAT pathway (Figure 9) (Kučerová et al. 2016). Another paper using transcriptional profiling of pupal diapause in the flesh fly (*S. crassipalis*) reported that 10 transcripts were differentially expressed by at least 2-fold during diapause across 3 different species, pupal diapause of the flesh fly, adult reproductive diapause of the fruit fly (*D. melanogaster*) and dauer formation of the nematode (*C. elegans*). These findings suggest that while some similarities exist for diapause signaling across species, it is likely that there are multiple means to a similar end (Ragland et al. 2010). More importantly, had this paper examined a different insect species that possesses an adult diapause, rather than pupal, as well as a later stage of diapause in C. elegans (i.e. ARD), more genes may have been found in common. From these gene expression studies, it is clear that an overlap exists between key developmental pathways governing lifespan and diapause.

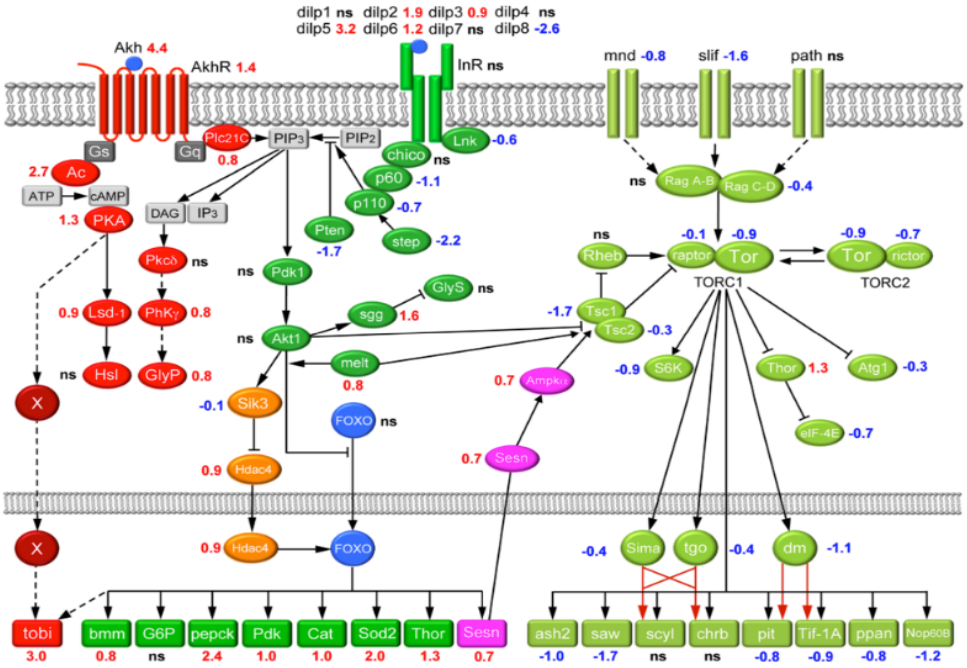

Figure 9 Overview of Gene Expression Changes During Diapause
Changes in gene expression levels for Akh, IIS and TOR signaling. Log fold changes are show in red for upregulated and blue for down regulated. Adapted from Kučerová et al (2016).

For diapause, the neuroendocrine axis and its associated organs have been studied in some detail in other insect species, such as the mosquito and the flesh fly. Work in these organisms has shown the involvement of the insulin producing cells (IPCs), corpora allata (CA), and the fat body (FB), in regulation of diapause initiation and its associated phenotypes (Xu et al. 2012). Ablation of the corpora allata in the mosquito lead to an increased incidence of diapause induction even under normal conditions (Spielman 1974). The corpora allata is the main production center for Juvenile Hormone (JH). During normal reproductive development, the CA is stimulated to produce JH by insulin signaling from the IPCs (Schiesari et al. 2011). This then leads to the production of Ecdysteroids which lead to the uptake of yolk proteins in the ovary (Figure 10). It has been demonstrated that topical application of JH to diapausing flies releases the hold on their reproductive development, but whether or not this also removes the delay in organismal senescence has yet to be examined (Saunders et al. 1990).

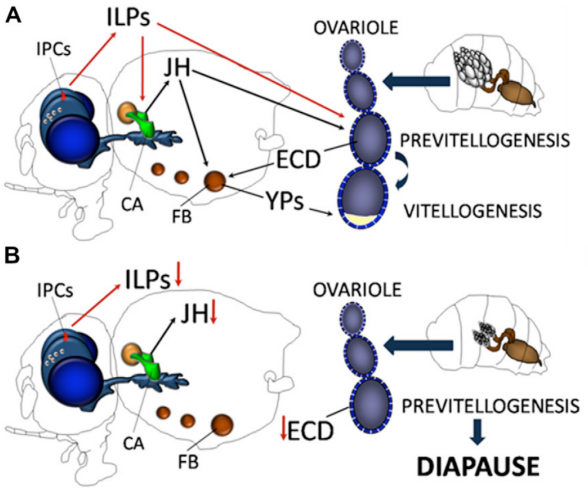

Figure 10 Endocrine Signaling During Diapause
During favorable conditions normal hormonal signaling occurs, governed by the IPCs. During diapause the presence of dilps decreases and prevents normal endocrine signaling resulting in delayed reproductive development. Adapted from Schiesari et al (2011).

As in the regulation of aging, the involvement of the insulin signaling pathway in diapause is great, and has been shown in multiple model organisms (Sim and Denlinger 2013; Sim and Denlinger 2008; Schiesari et al. 2016; Kimura et al. 1997; Hand et al. 2016). Not only does this signaling occur locally, but systemically as well, further indicating the role of hormonal control over diapause regulation. Overexpression of dILPs-2 and 5 were shown to induce vitellogenesis in diapausing flies back to normal levels (Schiesari et al. 2016).

Another paper using two fly strains with loss of function mutations in insulin like proteins (*dilp2/dilp3* & *dilp5*) found that these lines showed an increase in diapause induction (Kubrak et al. 2014). Implicating the role of specific dilps in regulating diapause induction and maintenance. Interestingly several papers have shown that levels of these transcripts increase during diapause (Kankare et al. 2016; Zhao et al. 2016; Kubrak et al. 2014; Kubrak et al. 2016), whether this is through increased transcript stability or increased expression is unclear, but this could allow for a quick recovery upon return to favorable conditions. Another implication of altered IIS signaling during diapause is that natural polymorphisms in the downstream target PI3K have been shown to affect the tendency of flies to enter diapause (Schiesari et al. 2011; Sim and Denlinger 2013; Kubrak et al. 2014). Other studies of diapause in *D. melanogaster* have shown that in addition to a decreased metabolic state and delayed senescence, the immune/stress response increases during this time. For more information on the heightened immune/stress response during diapause please see the following citations (Kubrak et al. 2014; Kubrak et al. 2016; Kučerová et al. 2016).

Traditionally *Drosophilid* diapause is most easily distinguished by observing the degree of egg chamber development in females. Flies are considered to have entered diapause if both ovaries fail to develop beyond stage 8 (Figure 11), presented as the proportion of pre-vitellogenic ovaries to the total number present within a group (Lumme et al. 1974; Saunders et al. 1989; Kubrak et al. 2014). This delayed development is most robustly seen at the 3 week diapause period, and afterward reproductive development resumes, with chorionated eggs seen after 6-8 weeks of diapause (Figure 12). Classification using this method indicates that the diapause trait within a population of *Drosophila* shows a wide range of variation, unlike other species where diapause induction within a given group

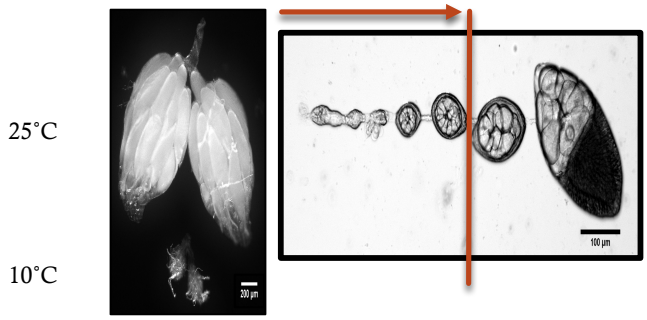

Figure 11 Control vs Diapause Reproductive Development
The difference between normal ovary development at 25˚C in comparison to diapause ovaries at 10˚C is easily seen. Termination of reproductive development at 10˚C occurs at pre-vitellogenic stages, indicated by the red line.

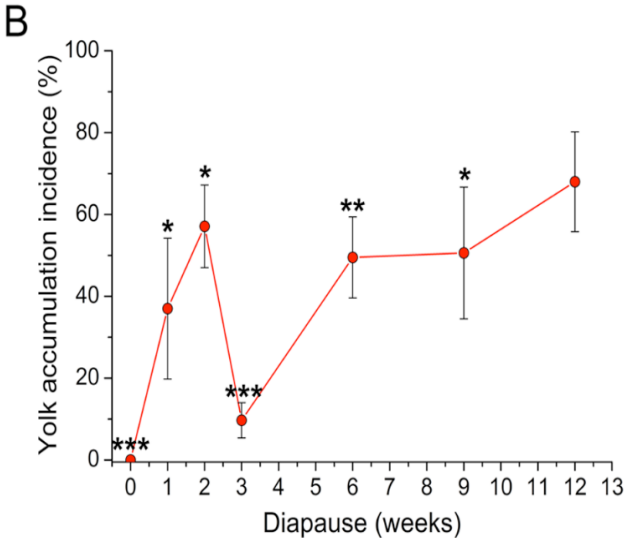

Figure 12 Percent Vitellogenesis Throughout Diapause
Yolk accumulation overtime during diapause. The lowest accumulation of yolk occurs at 3 weeks of diapause induction, and reproductive development appears to recommence onward. Adapted from Kubrak et al (2014).

is usually a binary decision. The existence of this complex trait on a continuous scale provides an excellent means with which to apply the use of Genome-wide Association Studies (GWAS) to identify potential genomic variants associated with this trait. Since the pause in reproductive development does not persist throughout diapause, and heavily weights the definition of diapause on reproductive component, this study utilizes the changes in lifespan and fecundity seen to determine and quantify diapause induction for use in GWAS.

Genome-Wide Association Studies: The Drosophila Genetic Reference Panel

As diapause involves several complex traits, it comes as no surprise that it itself, is a complex trait. A complex trait, in the most general sense, is a trait that does not follow simple patterns of Mendelian inheritance, involves the influence of multiple genetic variants within multiple genes (polygenic), and has a range of phenotypes (continuous). The integration and summation of multiple distinct signaling inputs of the resultant trait prove challenging to

study, due to the difficulty of untangling them with typical reductionist molecular and biochemical methods. Initial attempts at identifying components of such complex phenotypes has involved the use of trait mapping, specifically, quantitative trait loci (QTL) mapping, with little success. While these methods of genotype-phenotype mapping can result in the discovery of crucial genes with large weights on the influence of the phenotypic output, they are poorly powered for use with common genetic variants with lesser influence on the phenotype. Recent developments in sequencing capabilities and statistical genetics, have created improved methods of association mapping, allowing the identification of more common alleles with effects on the phenotype. Therefore Genome-wide association studies (GWAS) provide a unique and promising way to examine complex traits such as diapause.

Linkage Disequilibrium: QTL Mapping and GWAS Logistics

GWAS take their rudimentary shape from their historical kin, QTL mapping. These methods employ the fact that some alleles at different genetic loci have non-random, or dependent, association within a population, also known as linkage disequilibrium. Linkage disequilibrium is affected by factors such as selection, rate of recombination, rate of mutation, genetic drift, population structure, and physical genetic linkage. While both GWAS and QTL mapping utilize the concept of linkage disequilibrium, QTL studies measure fewer markers within a closely related cohort (familial based), whereas GWA studies assess more markers within a larger population.

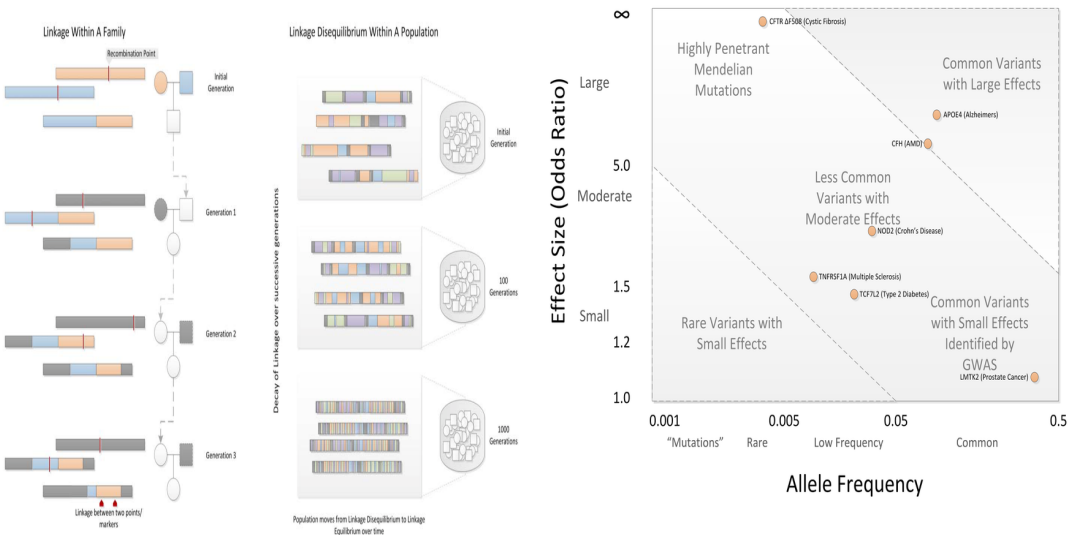

Figure 13 Visualization of Linkage, Linkage Disequilibrium & Effect Size
Linkage analysis allows for detection of genetic variants with strong/high effects on the resulting phenotype. As alleles become more varied within the population, and fall out of equilibrium, more powerful tests are need to detect causal variants. Figure Adapted from Bush & Moore (2012).

While linkage analysis has contributed to critical discoveries of genes involved in rare diseases, such as mutations in the *CFTR* gene that cause cystic fibrosis, and the DNA repeats that cause Huntington's Chorea, they are poorly powered to detect more common genetic variants with low phenotypic penetrance (Figure 13). This has led to the hypothesis that multiple alleles, each with a small phenotypic effect, sum up to cause a disease or phenotype (Bush and Moore 2012). Rapid advances in next-generation sequencing, which allows for high resolution mapping of genomic architecture, has overcome many of the challenges for QTL studies. Along with advances in statistical population genetics, the ability to test the association of millions of variants with complex traits is now feasible.

GWAS Pros & Cons: Types & Their Application

In general, a GWAS utilizes genome sequencing of a population, along with association mapping techniques, and evolving statistical algorithms, to identify genomic variants (SNPs, indels, inversions etc.) associated to phenotypes. The outcome of GWAS studies is greatly

influenced by four main factors: population structure, effect size of trait, variant heritability, and sample size. The inherent genetic architecture of a population (population structure) can differ among groups within a population, leading to what is known as cryptic relatedness, creating subpopulations. This in turn can lead to false associations due to variants being present in one subpopulation and not the other due to natural genetic segregation. The second factor that influences GWAS is the variation of the phenotype itself among the population. Continuous traits with very subtle differences between individuals are difficult to detect in genome-wide association studies, therefore the larger the variation of the trait, the better the association testing. Heritability of the trait also influences association analysis. Heritability is a statistic that estimates the amount of variation in a phenotypic trait due to genetic variation between individuals, any remainder is generally assumed to be caused by environmental effects, or an unidentified additional causal variant. Finally, the sample size of the population greatly influences the outcome of GWAS, since the number of variants in common between individuals decreases as the number of genomes compared increases (Bush and Moore 2012; Stranger et al. 2011; Klein et al. 2012). Measuring many individuals within a population decreases the chances of variant association by random chance, by increasing the accuracy of the sampling distribution representation.

Two main categories exist for GWAS design, the case-control method, and the quantitative trait method. The case-control setup is the most common approach used in human studies of disease, where individuals are grouped into two groups, those with the disease phenotype (case) and those without (control). The genotypes of the two cohorts are compared and variants present in the disease group but not the control group are tested for significant association. The second, quantitative trait based method, is used in situations

where the disease or trait exists on a continuous scale. This method of analysis is best used for complex traits that likely have multiple variants with smaller effects on the phenotype working together to produce the overall outcome. One key factor is the definition of the phenotype in question and its quantification. The methods of measurement, and the heterogeneity of the resulting phenotypic distribution, greatly affect the power of a GWAS. Another issue found in many GWAS utilizing human subjects, is that of genetic mapping and genome architecture. Due to the costs of sequencing, many studies use incomplete genomes, which greatly reduces the power of such a study, causing the mapping of variants to be very difficult. In the case of the model organism *Drosophila melanogaster*, the small size of the genome allows for complete genome sequencing, eliminating this concern.

GWAS and the DGRP

Presently there exists a fully-sequence, inbred population of flies dubbed the Drosophila Genetic Reference Panel (DGRP). The DGRP was created in 2012 by Trudy Mackay and her team, as a publicly available resource for population genomics and quantitative trait analysis (http://dgrp2.gnets.ncsu.edu/). The DGRP was created using a natural population of *Drosophila* taken from Raleigh, North Carolina. This natural population was then subjected to 20 rounds of full-sibling inbreeding achieving near homozygosity, resulting in a population of 201 individual fly lines. The DGRP has been thoroughly characterized for recombination rates, Wolbachia (a *Drosophila* reproductive parasite) infection status, variations in genome size, molecular population genetics for indels and inversion, and functional analyses of segregating genomic variants, providing a high-resolution detailed map of the genetic architecture for the population. This platform allows users to input quantitative phenotypic data and performs linear regression to determine

associations. These 201 lines possess large measurable genetic variation across multiple traits, lack strong underlying population structure, and have shown rapid decay in linkage disequilibrium making them ideal for performing GWAS (Mackay and Huang 2018; Mackay et al. 2012; Ellis et al. 2014; Huang et al. 2014).

Several labs have already implemented this tool to study a variety of traits. One group examined the complex trait of longevity and as proof of concept, found genes previously implicated in lifespan in addition to several gene candidates previously unassociated with lifespan (Ivanov et al. 2015). Other studies have examined traits such as: cold resistance, chill coma recovery, starvation, stress resistance, lifespan and fecundity, as well as several others, for which the data are available on the DGRP website. As these studies rapidly produce large amounts of data, I strongly agree with a statement made in a review of GWAS (McCarthy et al. 2008), that these studies and their likeness should first and foremost be used as hypothesis-driving guides. Enabling researchers to prioritize and focus studies towards determining genetic drivers of diseases and other complex traits, such as diapause. In this thesis, these tools are utilized to elucidate novel candidate genes involved in the diapause process, which could one day prove a radical system for studying the underlying contributors to the degenerative process known as aging.

Chapter II: GWAS of Diapause in the Drosophila Genetic Reference Panel

Introduction

The DGRP platform allows the user to input sample means (for each individual DGRP line), where regression algorithms are performed and then adjusted for potential covariates, such as Wolbachia infection, and 5 major inversions (In_2L_t, In_2R_NS, In_3R_P, In_3R_K, In_3R_Mo). This accounts for population structure or cryptic polygenic relatedness, which can result in false positive associations. The resulting analyses return a list of associated SNPs along with information regarding their location on the chromosome (intergenic region or genetic region), their nature (insertion, deletion etc.), their potential effect on genes (missense, nonsense, start gained etc.), associated gene annotations, and transcription factor regulation annotations. In the work described here, diapause is defined by the following criteria; the ability to survive diapause conditions, be capable of reproduction post-diapause, and maintain a relatively normal lifespan post-diapause. Multiple methods of normalizing and inputting the diapause lifespan and fecundity data were examined, before ultimately deciding to utilize the change in diapause lifespan and fecundity for data input into the GWAS, as the distribution of this data satisfied the criteria for Type III ANOVA testing. Using the difference between control versus diapause lifespan and fecundity means to quantify this trait, two separate genome-wide association studies were performed. The resulting gene annotations for the top associated SNPs analyzed for category enrichments both individually, and in combination, and gene networks created for each.

Materials and Methods

Experimental Overview

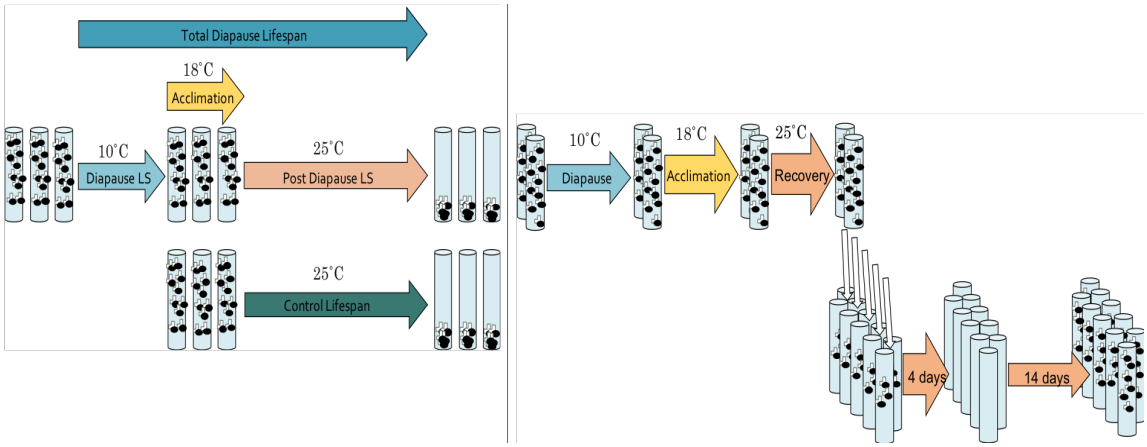

Figure 14 Overview of Experimental Designs
A) Virgin female flies were kept at 10˚C for 35 days, moved to 18˚C for a day to allow slow acclimation, before living out the remainder of their lives at 25˚C. For controls, virgin females were kept at 25˚C for the entirety of their lifespans. B) Virgin female flies were kept at 10˚C for 35 days, moved to 18˚C for a day allowing slow acclimation, and then moved to 25˚C for a day before mating to two Canton S male flies. Parental flies were mated for 4 days before removal, and progeny counted 14 days after.

Diapause Induction: Fecundity & Lifespan Measurement

For diapause induction, female virgin flies were maintained on standard yellow food, in vials at a density of 10 flies per vial for longevity and fecundity studies, and kept in a climate chamber at 10˚C under a light cycle of 8L:16D (light: dark). Controls were maintained at the same density at 25˚C, under a light cycle of 12L:12D. Lifespan was measured using 3 replicates of 10 flies for each line, that were subjected to winter like conditions for 35 days (5 weeks), then removed and slowly acclimated back to room temperature by spending 1 day at 18˚C, before moving to 25˚C for the remainder of their lives (Figure 14A). Vials were checked daily and fly day of death recorded. Vials were flipped as needed, or every 3 days. Some flies were lost during flipping and discounted from further analysis. Fecundity was measured using virgin females for each strain, mated to 2

Canton S (CS) males on the third day of post-diapause development. A sample size of 20 total individual flies was attempted for fecundity measurements. Flies were allowed to mate for 4 days before adults were removed. Progeny were then counted 14 days after mating in order to detect the maximum number of progeny from one generation (Figure 14B). Diapause was then quantified using the difference between the control and diapause means, by submitting the control and diapause data into the DGRP GWAS pipeline under "male" and "female" categories respectively.

Population & Line Statistical Analyses

All data were plotted using ggplot2 in the tidyverse package in R. All statistics were computed using base R and additional packages. Individual line data for the DGRP, CS, and w^{1118}, can be found in the Supplemental Material. Wilcoxon-rank sums tests were performed to determine if any significant changes occurred between the diapause and control data for both lifespan and fecundity, and the resulting data for individual line analysis can be found in the Supplemental Material.

Genome-wide Association Studies

Genomic variants were filtered based on the minor allele being present in at least four of the DGRP lines, resulting in 2,490,165 SNPs and 77,756 microsatellites used for association with the diapause phenotype, using Bonferroni corrections for multiple testing. Using the difference between the control and diapause lifespan, as well as the control and diapause fecundity, as the quantifiers for two separate GWAS of diapause. Means for 195 DGRP lines were measured for the fecundity analysis, and 190 DGRP lines for the lifespan analysis. The discrepancy in the number of DGRP lines utilized is due to the inability to

collect data for both categories under both conditions (control and diapause) for some of the DGRP lines. All Genome-wide association studies were carried out using the online platform designed and developed by Trudy Mckay et al (Huang et al. 2014; Mackay et al. 2012). Diapause means were entered under the "Male" category of the platform, while the control data was entered under the "Female", in order to test for variants associated to the difference between the two conditions. Results from the two GWAS were given as depicted in Figure 15 Four methods of regression analysis are performed for the following categories: female (control), male (experimental), average of the two, and the difference between the two. P-values are given, for the simple linear regression models ("XPval") and for the mixed effects models ("XMixedPval").

	ID	MinorAllele	MajorAllele	RefAllele	MAF	MinorAlleleCount	MajorAlleleCount	FemaleEff	FemalePval	FemaleMixedPval	MaleEff	MalePval	MaleMixedPval	AvgEff	AvgPval	AvgMixedPval	DiffEff	DiffPval	DiffMixedPval	GeneAnnotation	RegulationAnnotation
1	X_13068569_SNP	A	G	G	0.1695	30	147	1.021	0.2415	0.238511223	6.345	5.02E-09	5.76E-09	3.683	4.72E-06	3.74E-06	-5.324	2.69E-06	4.23E-06	SiteClass[FBgn0085	TF_binding_site|BDTI
2	X_13067870_SNP	G	T	T	0.1667	30	150	1.278	0.1422	0.138831471	6.041	4.39E-08	3.20E-08	3.659	5.65E-06	4.18E-06	-4.763	4.26E-05	4.16E-05	SiteClass[FBgn0085	TF_binding_site|BDTI
3	X_7974208_SNP	C	A	A	0.0582	11	178	0.4339	0.7493	0.747035253	9.346	3.94E-08	3.67E-08	4.89	9.34E-05	8.50E-05	-8.913	6.66E-07	6.08E-07	SiteClass[FBgn0029	TF_binding_site|BDTI
4	3L_1111454_SNP	C	G	G	0.07487	14	173	-1.727	0.155	0.150001603	6.895	8.61E-06	7.38E-06	2.584	0.02344	0.021965286	-8.622	6.84E-08	5.69E-08	SiteClass[||||,Tran	TF_binding_site|BDTI
5	2L_2855787_INS	ATGGA	A	A	0.1421	26	157	-1.075	0.2498	0.239981458	5.481	3.87E-06	2.76E-06	2.203	0.01157	0.01023244	-6.557	1.03E-07	6.09E-08	SiteClass[FBgn0031	TF_binding_site|BDTI
6	2L_2855805_SNP	A	C	C	0.1429	26	156	-0.8362	0.3724	0.361460376	5.673	1.83E-06	1.17E-06	2.419	0.005742	0.004758603	-6.509	1.27E-07	7.82E-08	SiteClass[FBgn0031	TF_binding_site|BDTI
7	3L_1412443_SNP	A	G	G	0.2151	40	146	0.7237	0.3555	0.348507349	5.243	1.07E-07	8.01E-08	2.983	3.60E-05	2.85E-05	-4.519	1.49E-05	1.24E-05	SiteClass[FBgn0003	-
8	2R_19509740_SNP	G	C	C	0.3135	58	127	1.787	0.009625	0.008603753	-3.042	0.0007208	0.000602867	-0.6276	0.3405	0.331913448	4.829	1.27E-07	1.02E-07	SiteClass[||||,Tran	TF_binding_site|BDTI
9	2L_2855816_SNP	C	T	T	0.1374	25	157	-1.292	0.1729	0.164749923	5.246	1.49E-05	1.10E-05	1.977	0.02601	0.023739349	-6.538	1.95E-07	1.15E-07	SiteClass[FBgn0031	TF_binding_site|BDTI
10	2L_13768860_SNP	G	A	A	0.1093	20	163	5.154	2.78E-07	1.95E-07	1.951	0.152	0.146264626	3.553	0.0002181	0.000186671	3.204	0.02452	0.021888838	SiteClass[||||,Tran	-
11	3L_9059114_SNP	A	G	G	0.06383	12	176	4.266	0.0008692	0.000804819	7.855	2.36E-06	1.82E-06	6.061	3.43E-07	2.68E-07	-3.589	0.0438	0.041720426	SiteClass[FBgn0000	TF_binding_site|mE1
12	2R_19509741_SNP	C	T	T	0.3017	54	125	1.547	0.02853	0.027291203	-3.218	0.0005387	0.000408801	-0.8355	0.2172	0.208367739	4.764	4.51E-07	3.49E-07	SiteClass[||||,Tran	TF_binding_site|BDTI
13	2R_19751690_SNP	C	G	G	0.05319	10	178	-2.733	0.0536	0.050987924	6.744	0.0002378	0.000200736	2.005	0.1344	0.129905237	-9.476	5.15E-07	3.84E-07	SiteClass[FBgn0025	TF_binding_site|BDTI
14	3R_4610617_SNP	T	G	G	0.08696	16	168	-0.673	0.5595	0.552308627	6.942	2.11E-06	1.49E-06	3.135	0.003598	0.003021925	-7.615	5.42E-07	4.04E-07	SiteClass[FBgn0263	TF_binding_site|mE1
15	X_13067738_SNP	C	T	T	0.1793	33	151	1.086	0.1959	0.19068804	5.347	6.24E-07	4.15E-07	3.216	3.54E-05	2.78E-05	-4.26	0.0001808	0.000142468	SiteClass[FBgn0085	TF_binding_site|BDTI
16	X_13068820_SNP	A	G	G	0.3526	61	112	0.01706	0.9801	0.980233206	4.412	5.16E-07	4.58E-07	2.215	0.0005034	0.000542221	-4.395	2.07E-06	1.66E-06	SiteClass[FBgn0085	TF_binding_site|BDTI
17	3L_1412385_SNP	C	T	T	0.3081	57	128	0.5765	0.4069	0.402710208	4.409	4.26E-07	4.77E-07	2.493	0.00011	9.33E-05	-3.832	2.46E-05	3.42E-05	SiteClass[FBgn0003	-
18	X_13067739_SNP	A	G	G	0.1749	32	151	0.9887	0.2456	0.239842354	5.341	9.04E-07	6.08E-07	3.165	6.00E-05	4.78E-05	-4.353	0.0001592	0.000123637	SiteClass[FBgn0085	TF_binding_site|BDTI
19	X_16726107_SNP	T	A	A	0.2888	54	133	1.308	0.06086	0.0599477	4.423	6.71E-07	6.32E-07	2.865	8.13E-06	8.22E-06	-3.114	0.001089	0.000981061	SiteClass[FBgn0030	TF_binding_site|BDTI
20	3L_7436493_SNP	G	C	C	0.06452	12	174	5.306	2.29E-05	2.66E-05	6.454	0.0001273	0.000104007	5.88	6.74E-07	6.41E-07	-1.149	0.5215	0.516827027	SiteClass[FBgn0259	-
21	X_11348527_SNP	T	A	T	0.371	69	117	1.525	0.02072	0.019482719	-2.818	0.001026	0.000917601	-0.6469	0.3021	0.296366965	4.343	7.47E-07	6.53E-07	SiteClass[FBgn0262	TF_binding_site|BDTI
22	3L_11508729_DEL	C	CCC	CCC	0.05714	10	165	-2.98	0.03426	0.033522541	6.324	0.0007141	0.000517735	1.672	0.2126	0.208100008	-9.304	1.53E-06	6.80E-07	SiteClass[FBgn0036	-
23	X_13068502_SNP	C	G	G	0.1858	34	149	0.8192	0.3114	0.318902011	5.156	8.34E-07	8.35E-07	2.988	6.26E-05	8.70E-05	-4.337	0.0001207	8.87E-05	SiteClass[FBgn0085	TF_binding_site|BDTI
24	3R_13509752_SNP	C	T	T	0.08556	16	171	5.402	1.07E-06	8.44E-07	3.582	0.01597	0.015117102	4.492	1.88E-05	1.65E-05	1.82	0.2461	0.240218127	SiteClass[FBgn0053	TF_binding_site|mE1

Figure 15 Visualization of DGRP GWAS Output
Results provide the ID of the variant associated, the respective minor, major and reference allele, along with counts for each alleles' presence. P-values are given for the 4 different categories (female, male, average and difference) and for each of the methods of regression (single or mixed), resulting in 8 total p-values per variant. The effects of each variant on the phenotype are also given for each of the 4 categories.

Gene Network & GO Analyses

Gene sets from the two analyses were examined for biological component, molecular function, and pathway enrichment, using the ClueGo (v2.5.2) plug-in for Cytoscape (v3.6.1).

Enrichment was tested using right-sided hypergeometric tests (overrepresentations) with Bonferroni step down corrections applied.

Results

Lifespan & Fecundity Statistics

To gain insight on the variation of lifespan and fecundity traits amongst the DGRP population, the average normal (control) lifespan and average normal (control) fecundity were examined. 193 strains were used (including CS and w^{1118}) resulting in 6,239 flies observed for the distribution of the population's control lifespan. The average control lifespan followed a very normal distribution (skew: -0.59), and ranged from 2 to 82 days, with the mean and median lifespans being 44.29, and 46 days respectively (sd: 16.47 days; SEM: 0.203). For the typical fecundity population distribution, 196 strains, including CS, were observed totaling 2,598 flies. The overall population distribution for this trait showed a weighted distribution towards lower progeny values (skew: 0.63), with a range of 0-115 total progeny produced from the four-day mating and egg laying period. The mean fecundity for the population was found to be 30.42 progeny, with a median of 27 (sd: 23.48 flies; SEM: 0.461). Lifespan and fecundity were then examined under diapause conditions as previously described.

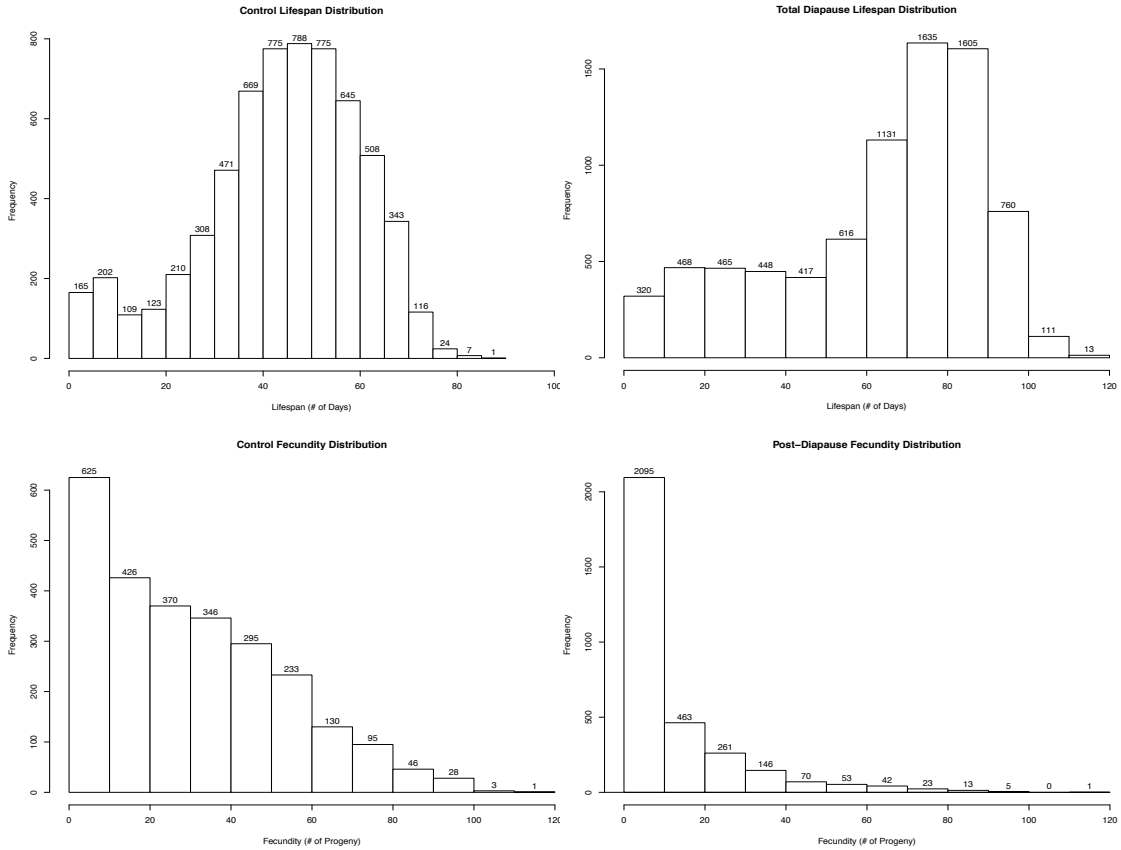

Figure 16 Population Distributions for Control & Diapause Lifespan & Fecundity
Control population distributions for lifespan and fecundity are presented on the left, while diapause population distributions are presented on the right. Numbers above bars represent the number of observations within that bin.

For diapause lifespan, a total of 200 strains were examined (including CS and w^{1118}), with observations of 7,989 flies. The total-diapause lifespan (from eclosion until death; TDLS) ranged from 2 to 115 days, with a mean and median TDLS of 63.90 and 71 days respectively. While the control lifespan distribution of the DGRP population follows a normal distribution, the distribution of the average total diapause lifespan, though relatively normal (skew: -0.77), clearly shows a shift towards higher means (Figure 16). This increase

in the total average lifespan of the population, from about 44 days to almost 64 days was significant ($p = < 2.2\ e^{-6}$). For diapause quantification using fecundity, the post-diapause fecundity was observed after 2 days of recovery from diapause, for 195 lines (including w^{1118}), totaling 3,172 flies. The post-diapause (PD) fecundity ranged from 0-113 progeny, with a mean and median progeny production of 11.23 and 4 (sd: 16.41; SEM: 0.291). The general distribution for the fecundity after diapause showed a significant shift towards lower values of progeny produced (skew: 5.15), with an overall decrease in average fecundity from 30.43 to 11.23 ($p = < 2.2\ e^{-6}$). A statistical summary can be found for the control and diapause populations in Table 1.

Table 1 Summary of Control & Diapause Lifespan & Fecundity Population Statistics

Trait	n	Mean	SD	SE	Med	Trimmed	Adjusted SD	Min	Max	Skew	Kurtosis	25th Quant	75th Quant	Wilcoxon Results	
														W	p-value
Control Lifespan	6239	44.29	16.5	0.2	46	45.51	14.83	2	86	-0.591	-0.017	35	56	37723000	<2.2e-6
Diapause Lifespan	7989	63.9	25.8	0.3	71	66.34	22.24	2	115	-0.771	-0.448	48	83		
Control Fecundity	2598	30.43	23.5	0.5	27	28.41	25.2	0	115	0.635	-0.285	11	46	1907400	<2.2e-6
Diapause Fecundity	3172	11.23	16.4	0.3	4	7.633	5.93	0	113	2.159	5.154	0	16		

While majority of the DGRP lines met, or surpassed their control lifespan based on their total diapause lifespan as depicted in Figure 17, most did not meet or surpass their normal lifespan based solely on their post-diapause lifespan (Figure 18). 173 lines had a significant increase in their total lifespan when their total diapause lifespan was compared to controls, while 5 lines showed a significant decrease in overall lifespan when total-diapause lifespan was compared to controls, based on Wilcoxon rank-sum tests. The remaining lines showed no significant difference in their lifespan change (Figure 17). When comparing the post-diapause lifespan to that of controls, the majority (153 lines) showed an overall

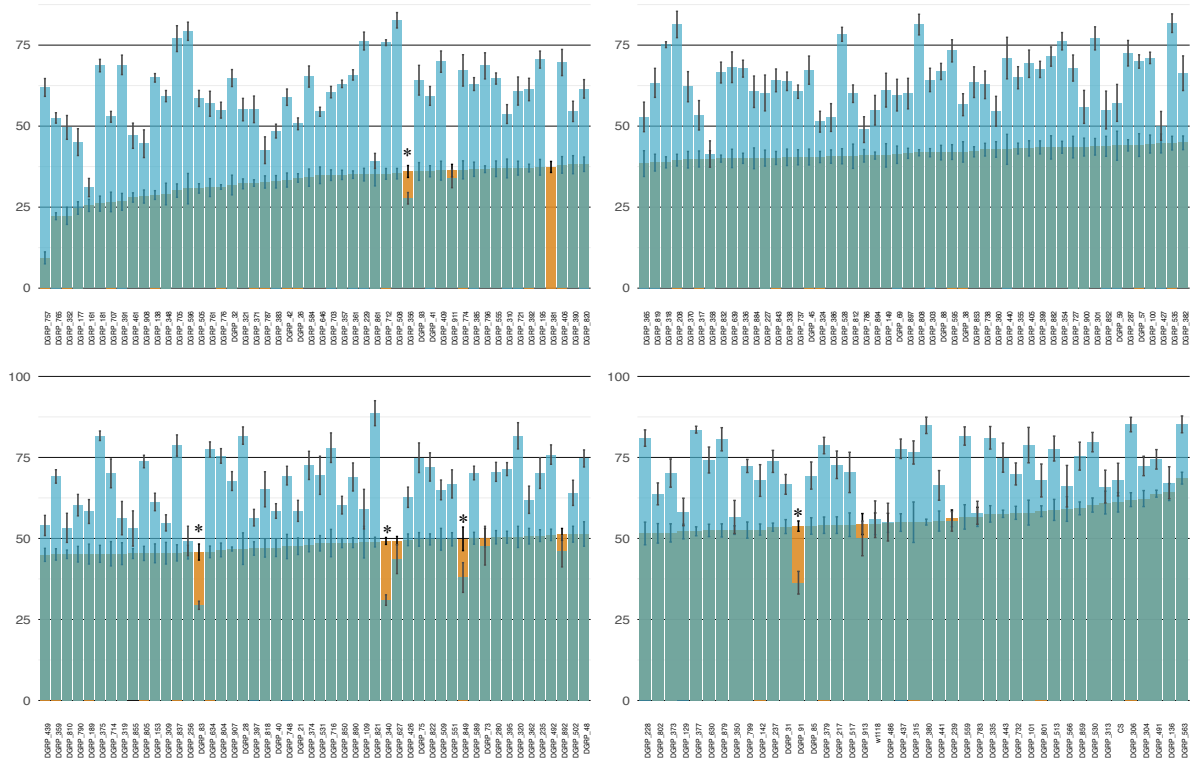

Figure 17 Bar Plot of Average Control & Total Diapause Lifespans
Bar plot depicts the mean control lifespans (orange) in increasing order, with the corresponding mean total diapause lifespans (blue). Asterisks demarcate total diapause lifespans that showed a significant decrease (p-value < 0.05) compared to controls based on one-sided Wilcoxon Rank-Sums tests.

significant decrease from their normal lifespans. A handful of lines (23) showed no significant change in their PDLS compared to controls, essentially doubling their overall lifespan. Remarkably, 14 lines showed a significant increase in their lifespan post-diapause (Figure 18). Table 2 summarizes the results for select DGRP lines that showed significant interesting lifespan shifts. In Figure 20 most of the DGRP strains showed a significant decrease in their fecundity capacity post-diapause (157 lines), and several lines failed to produce any progeny despite survival from diapause conditions (11 lines). Intriguingly, four of the strains showed a significant increase in their PD fecundity going from means ranging from 0-56 for control fecundity to 0-65 post-diapause. Table 2 summarizes the changes seen for these four DGRP lines, with associated p-values. Results from the Wilcoxon Tests

between all the individual DGRP lines for lifespan and fecundity can be found in the Supplemental Material available upon request.

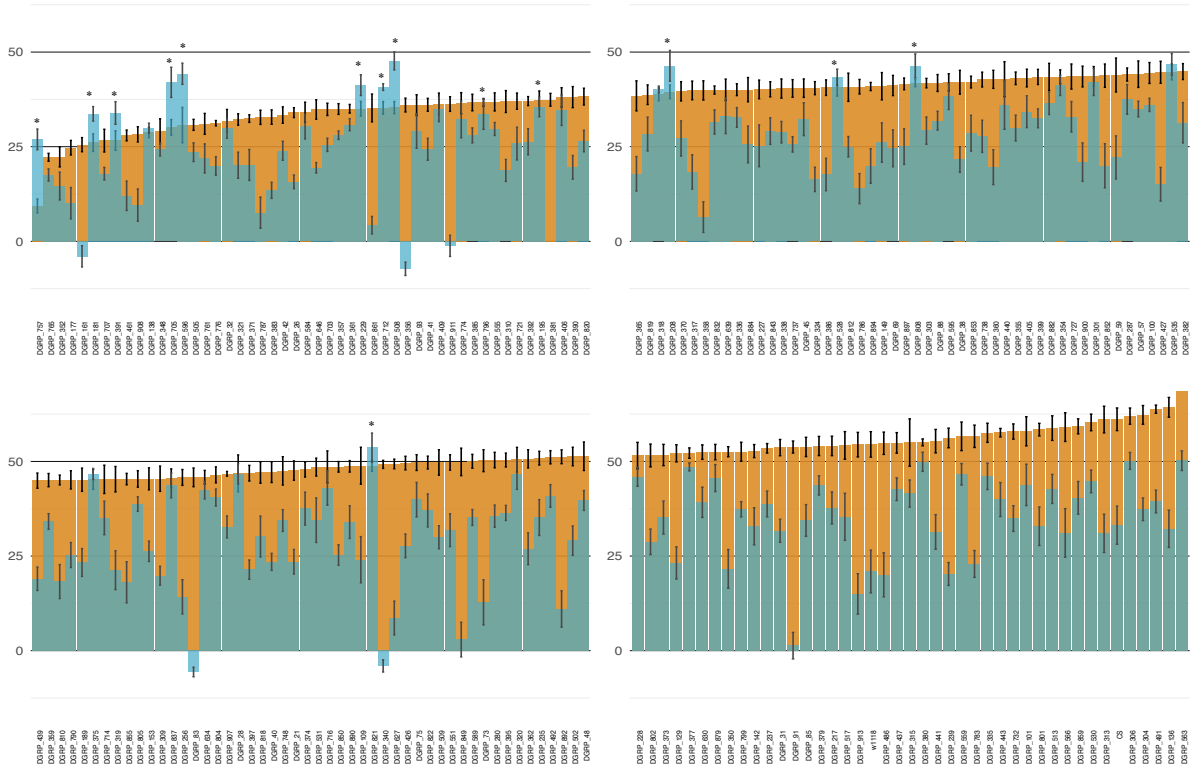

Figure 18 Bar Plot of Average Control & Post-Diapause Lifespans
Bar plot depicts the mean control lifespans (orange) in increasing order, with the corresponding mean post-diapause lifespans (blue). Asterisks demarcate post-diapause lifespans that showed a significant increase (p-value <0.05) compared to controls based on one-sided Wilcoxon Rank-Sums tests.

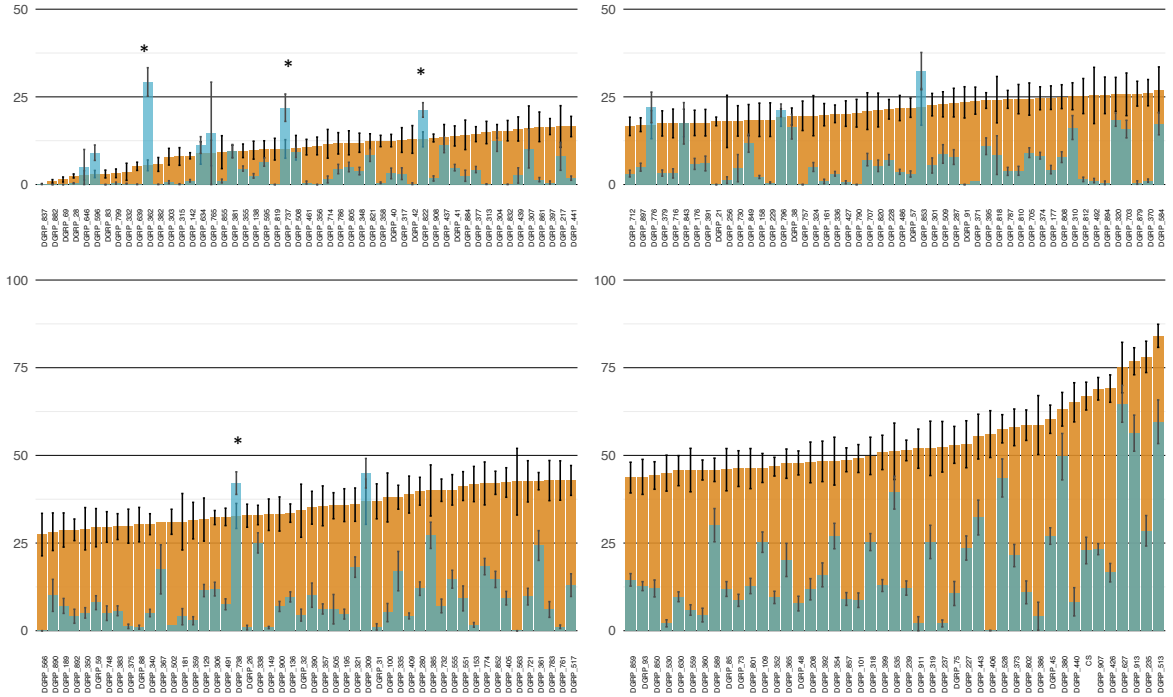

Figure 19 Bar Plot of Average Control & Post-Diapause Fecundity
Bar plot depicts the mean control fecundity (orange) in increasing order, with the corresponding mean post-diapause fecundity (blue). Asterisks demarcate post-diapause fecundity that showed a significant increase (p-value <0.05) compared to controls based on one-sided Wilcoxon Rank-Sums tests.

Table 2 Summary of Interesting DGRP Lifespan & Fecundity Changes

Significant Decrease in Lifespan			Total Diapause Lifespan Statistics									Control Lifespan Statistics								
Strain	p-value	W	n	mean	sd	median	trimmed	mad	min	max	se	n	mean	sd	median	trimmed	mad	min	max	se
DGRP_340	2.42E-13	2193	40	30.98	10.21	33	31.44	6.67	3	50	1.614	59	49.22	8.19	49	50.12	5.93	7	61	1.066
DGRP_356	9.23E-07	1909.5	80	27.74	15.80	23	24.16	2.97	6	85	1.766	30	36.00	10.03	38	37.33	8.90	6	53	1.831
DGRP_83	1.17E-06	975	30	29.37	6.89	29	29.04	5.93	15	53	1.258	39	45.79	15.46	49	47.42	14.83	11	63	2.475
DGRP_849	0.026119902	764	30	37.93	25.09	47	38.04	28.91	2	74	4.581	40	49.88	22.99	50	49.97	30.39	8	86	3.634
DGRP_91	1.05E-05	738	30	36.33	19.18	30.5	33.00	11.12	16	85	3.502	30	53.80	8.69	51.5	53.79	8.90	35	70	1.587
Significant Increase in Lifespan			Post-Diapause Lifespan Statistics									Control Diapause Lifespan Statistics								
Strain	p-value	W	n	mean	sd	median	trimmed	mad	min	max	se	n	mean	sd	median	trimmed	mad	min	max	se
DGRP_181	0.000236856	1028	117	33.68	20.71	43	36.99	7.41	-27	56	1.915	30	26.07	12.73	26.5	26.29	14.08	3	52	2.324
DGRP_195	0.002304103	779	80	35.53	23.44	48	39.89	4.45	-33	55	2.621	30	37.30	13.27	40.5	39.58	2.22	6	52	2.422
DGRP_208	4.62E-08	89.5	30	46.37	22.26	53	53.00	2.22	-29	57	4.065	30	39.50	9.88	42.5	42.08	3.71	4	46	1.805
DGRP_229	0.000364814	193.5	27	41.26	14.29	45	43.13	7.41	5	58	2.750	30	35.03	10.64	38.5	37.13	6.67	3	47	1.943
DGRP_391	0.019962302	207.5	21	33.90	13.72	38	36.71	4.45	-23	43	2.994	30	26.63	14.20	34	27.75	8.15	3	42	2.592
DGRP_508	0.000319735	61.5	9	47.67	7.11	50	47.67	2.97	35	54	2.369	49	35.33	11.26	35	36.24	10.38	2	52	1.609
DGRP_528	0.027067716	319.5	30	43.33	11.80	48.5	46.17	2.97	1	53	2.155	30	40.70	12.77	41.5	41.33	9.64	5	63	2.331
DGRP_596	0.005128603	67.5	30	44.27	15.35	51	46.58	5.93	-8	61	2.802	10	30.70	14.80	33.5	31.63	6.67	4	50	4.681
DGRP_705	0.00021514	303	40	42.00	25.21	47	45.16	17.79	-25	77	3.986	30	30.10	11.27	32	30.88	9.64	2	51	2.058
DGRP_712	0.001479448	349.5	40	40.75	5.57	41	41.25	4.45	25	50	0.881	30	35.27	9.08	34	35.63	6.67	4	53	1.657
DGRP_757	1.64E-05	50.5	15	26.93	10.63	28	26.23	4.45	11	52	2.746	29	9.34	9.91	5	7.92	1.48	2	35	1.840
DGRP_796	7.50E-05	195.5	30	33.60	22.04	40	38.38	1.48	-27	67	4.024	30	36.70	5.66	37	37.63	1.48	8	41	1.034
DGRP_808	9.68E-05	342	50	46.34	22.45	54	51.10	7.41	-33	68	3.175	28	41.79	4.95	41	41.50	5.93	34	53	0.935
DGRP_821	0.000198994	301.5	40	53.65	24.35	63	57.72	10.38	-27	77	3.850	30	48.87	8.16	51.5	50.75	3.71	17	55	1.490
Significant Increase in Fecundity			Post-Diapause Fecundity Statistics									Control Fecundity Statistics								
Strain	p-value	W	n	mean	sd	median	trimmed	mad	min	mad	se	n	mean	se	median	trimmed	mad	min	max	se
DGRP_362	4.08E-05	40.5	20	29.25	18.20	30.5	29.06	24.46	0	65	4.070	17	5.47	6.33	2	5.00	2.97	0	18	1.534
DGRP_737	0.01058174	34.5	11	21.91	12.96	17	22.22	13.34	0	41	3.909	14	10.14	9.94	8	9.33	9.64	0	30	2.656
DGRP_738	0.026265248	30.5	11	42.09	10.68	42	42.00	11.86	25	60	3.221	11	32.73	11.84	26	31.33	4.45	22	56	3.570
DGRP_822	0.002876888	54	15	21.27	8.03	22	21.00	5.93	9	37	2.074	17	12.88	8.78	8	12.40	5.93	2	31	2.130

Genome Wide Association Studies of Diapause

Throughout the course of these experiments, several methods of quantifying and normalizing the lifespan and fecundity data were executed using the DGRP GWAS platform. Ultimately the difference between the control and diapause data was selected as the best method of examining the diapause phenotype. The results for this method are presented and discussed below, while results from previous methods of analysis are available upon request. Submitting the line means for diapause and control data under the “male” and “female” categories (respectively), results in an output containing top associated SNPs for multiple comparison tests performed on the difference between the two categories, each individually, as well as the average of the two.

For GWAS analysis of diapause versus control lifespan, the data was entered under the "male" and "female" categories and linear regression analyses performed. These analyses returned 122 total variants associated to the phenotypic data for all 8 of the test methods. 101 variants were associated to the difference between control and diapause lifespan for both the mixed and single models. 106 and 107 variants were associated to the diapause phenotype data alone for single and mixed models respectively. While control lifespan alone showed 56 variants associated to the phenotypic data for both single and mixed regression models. Manhattan plots depicting the distribution of SNPs associated to the difference between diapause and control lifespan as well as those associated with the diapause lifespan alone along the chromosomes can be seen in Figure 20. Table 3 contains a summary of all associated genes annotated for each of the regression analyses.

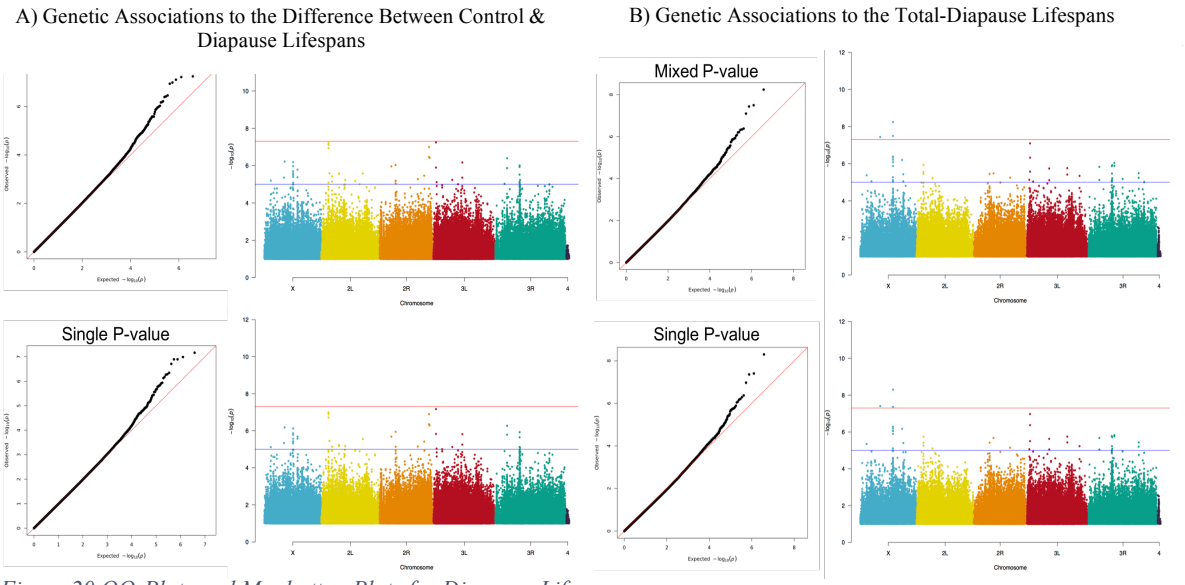

Figure 20 QQ-Plots and Manhattan Plots for Diapause Lifespan
A) GWAS Results based on the difference between Control and Diapause lifespan means. QQ-plots depict the correlation between the observed p-values (y-axis) and those expected to occur due to random association (x-axis). Manhattan plots show the variants associated per chromosome, filtered by those variants with p-values less than $1e^{-1}$. B) QQ-plots and Manhattan plots for the Total-diapause lifespan means alone. Both the Mixed p-value and Single p-value are shown.

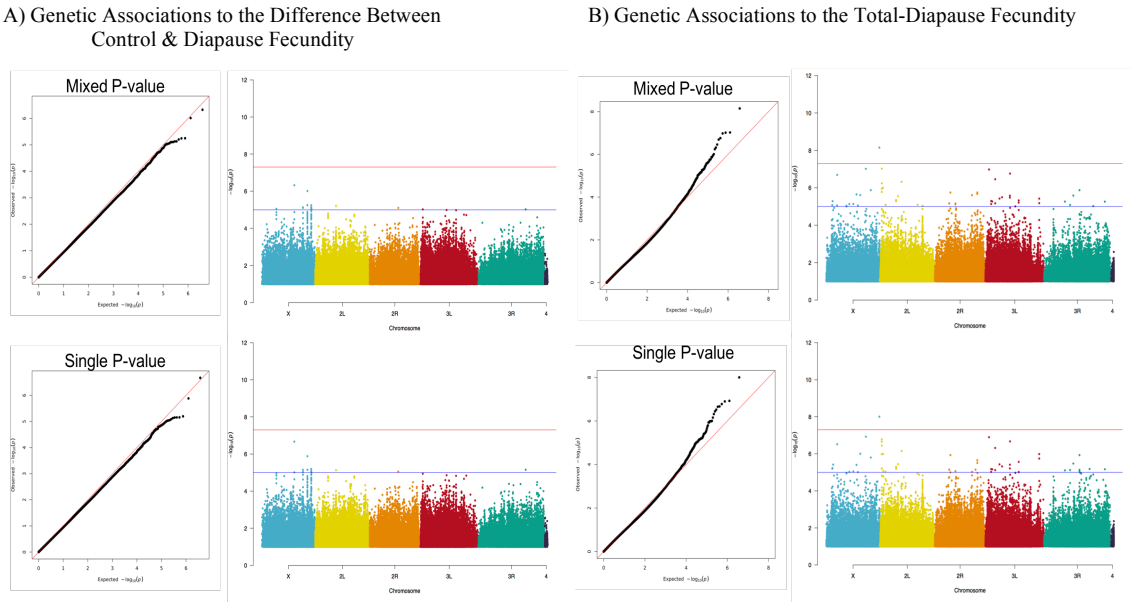

Figure 21 QQ-plots and Manhattan Plots for Diapause Fecundity
A) GWAS Results based on the difference between Control and Diapause fecundity means. QQ-plots depict the correlation between the observed p-values (y-axis) and those expected to occur due to random association (x-axis). Manhattan plots show the variants associated per chromosome, filtered by those variants with p-values less than $1e^{-1}$. B) QQ-plots and Manhattan plots for the Post-diapause fecundity means alone. Both the Mixed p-value and Single p-value are shown.

For GWAS analysis of post-diapause versus control fecundity, data were entered in the format described in the materials and methods section, and linear regression analyses performed. These analyses returned 107 total variants associated to the phenotypic data for all 8 of the test methods. 30 variants were associated to the difference between control and diapause lifespan for the single model, and 32 variants were found based on the mixed model. 91 and 90 variants were associated to the diapause phenotype data alone, for single and mixed models respectively. While control lifespan alone showed 99 variants associated to the phenotypic data for both single and mixed regression models. Table 3 summarizes the number of associated genes annotated for each of the regression analyses.

The corresponding quantile-quantile plots provide insight into any confounding variables that may be present during association analyses, such as underlying genomic population structure. These graphs plot the expected –log(p-values) vs the observed –log(p-

values), with the XY line representing a plot of SNPs associated by random chance. The typical shape of a "successful" GWAS should have an initial overlap with the XY line, that then curves away at lower p-values where "true" associations are more likely to exist. We can see that for both methods of lifespan analysis (both difference and diapause alone) return the stereotypical Q-Q plot observed for successful GWAS, with true associated SNPs falling above the XY-line, indicating a low likelihood of false positives. For the fecundity analysis, we see that the Q-Q plot for the diapause data alone produces satisfactory results, that indicate true associations at the indicated p-values (Figure 21B). However, for the associations to the difference between post-diapause fecundity and control fecundity we see a dip below the XY line, which indicates that for our most "significantly" associated variants display p-values that would've been obtained by random chance (Figure 21A).

Table 3 Summary of Diapause Lifespan & Fecundity GWAS Results

GWAS Results for Control & Total-Diapause Lifespan							
Method Used	**All**	**Difference**		**Diapause Data Alone**		**Control Data Alone**	
Regression Type	**Both**	**Single Pvalue**	**Mixed Pvalue**	**Single Pvalue**	**Mixed Pvalue**	**Single Pvalue**	**Mixed Pvalue**
# of Associated Variants	122	101	101	106	107	56	56
# of Gene Annotations	74	59	59	67	68	31	31
Intergenic Regions	26	20	20	20	20	25	25
GWAS Results for Control & Post-Diapause Fecundity							
Method Used	**All**	**Difference**		**Diapause Data Alone**		**Control Data Alone**	
Regression Type	**Both**	**Single Pvalue**	**Mixed Pvalue**	**Single Pvalue**	**Mixed Pvalue**	**Single Pvalue**	**Mixed Pvalue**
# of Associated Variants	107	30	32	91	90	99	99
# of Gene Annotations	77	22	25	68	68	63	63
Intergenic Regions	26	10	10	20	19	26	26

Gene Network Analysis & GO Enrichment Analysis

Using the 74 unique genes annotated in the results from the GWAS conducted using the difference between control and diapause lifespan, and the 77 unique genes annotated from the difference between diapause fecundity GWAS results, three gene network analyses

were conducted. The network generated from the lifespan related gene list (Figure 22) showed enrichment in several GO categories and one KEGG term. Five terms were found

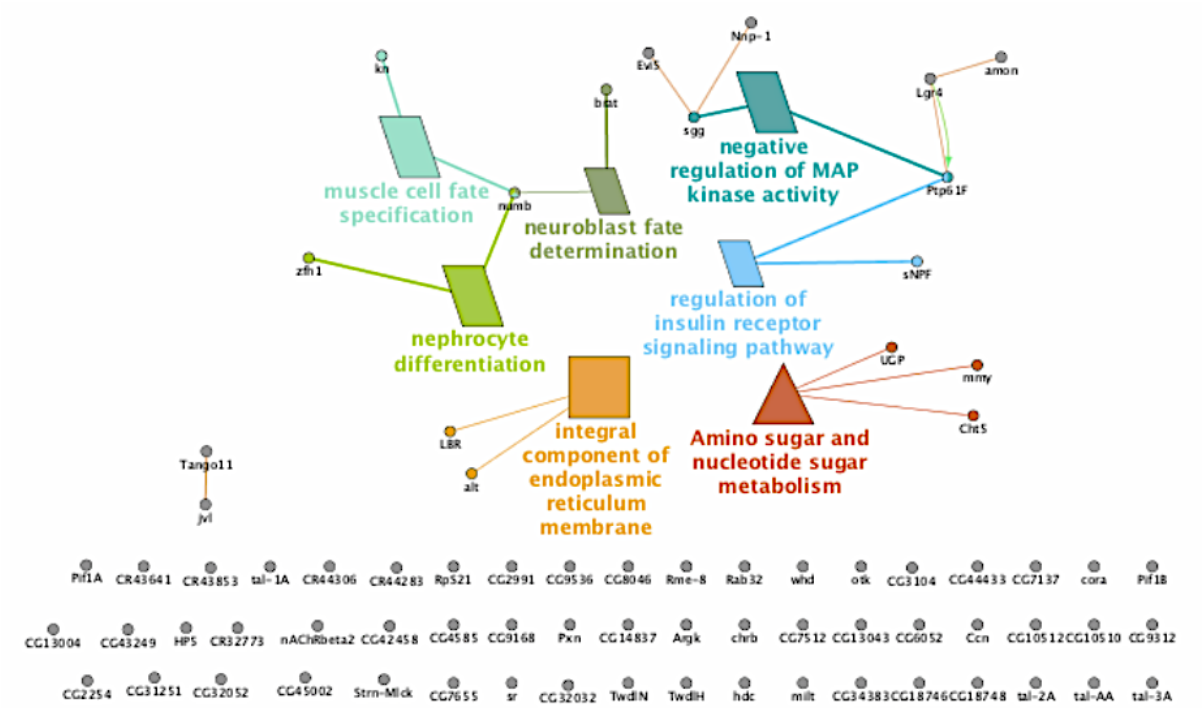

Figure 22 Gene Network of Control & Total Diapause Lifespan Genes
Gene network overrepresentation analysis for all genes annotated in the diapause fecundity based GWAS results. Circles are individual genes annotated in the data set, color coded by gene ontology involvement, grey circles are unassigned genes. Parallelograms show GO biological processes, squares show GO cellular components, and triangles depict KEGG pathways, represented in the gene set. Green arrows indicate positive gene regulation, while orange lines indicate protein-protein interactions.

significantly represented for GO biological process, one term for GO cellular component, and one for KEGG pathway. The gene network analysis produced from the fecundity related gene list (Figure 23) showed significant representation of six GO biological processes, two GO molecular functions, and three reactome pathways. An overview of all enrichment terms/categories is presented in Table 4 along with the genes contributing to each. Gene network analysis on the combined top associated genes for longevity and fecundity combined showed enrichment in several categories (Figure 24). While all categories that appeared in the separate network analyses were shown in the combined analysis, three novel reactome pathways emerged (Table 5).

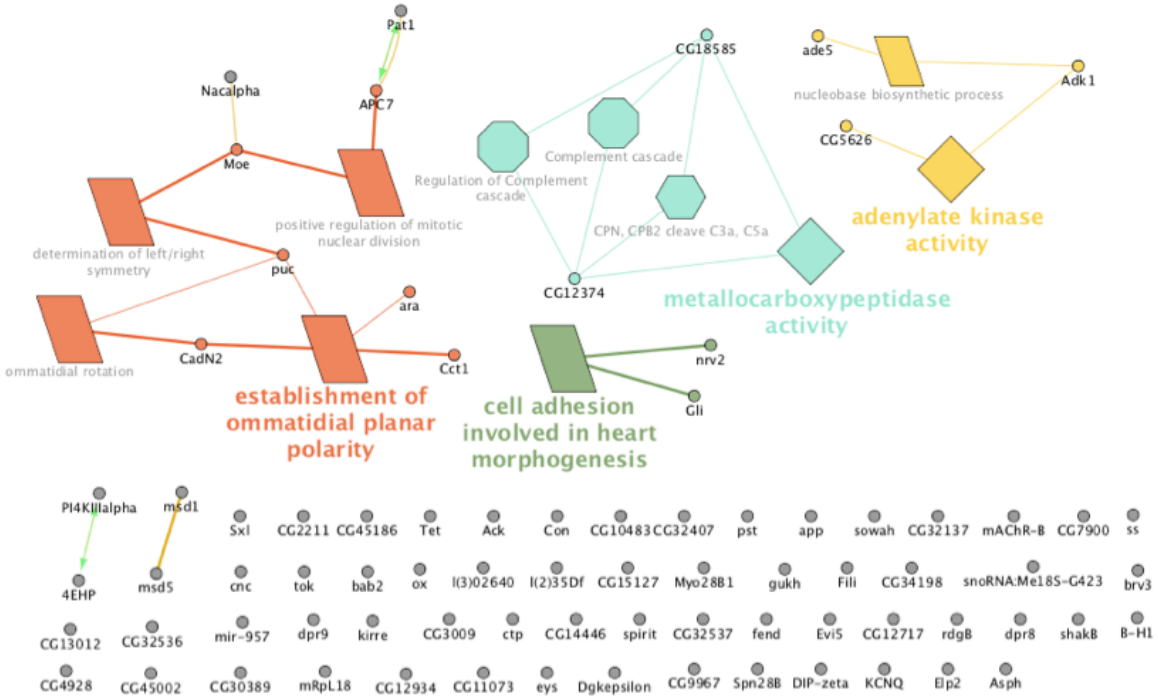

Figure 23 Gene Network Analysis of Control & Diapause Fecundity Genes
Gene network overrepresentation analysis for all genes annotated in the diapause fecundity based GWAS results. Circles are individual genes annotated in the data set, color coded by gene ontology involvement, grey circles are unassigned genes. Parallelograms show GO biological processes, GO diamonds show molecular functions, and octagons depict reactome pathways, represented in the gene set. Green arrows indicate positive gene regulation, while orange lines indicate protein-protein interactions.

Table 4 Summary of Diapause Lifespan & Fecundity Gene Network Results

Gene Network Results for Control & Diapause Lifespan				
GOID	GOTerm	% Associated Genes	Nr. Genes	Associated Genes Found
GO:0007400	neuroblast fate determination	7.14	2.00	[brat, numb]
GO:0030176	integral component of endoplasmic reticulum membrane	6.25	2.00	[LBR, alt]
GO:0042694	muscle cell fate specification	22.22	2.00	[kn, numb]
GO:0043407	negative regulation of MAP kinase activity	20.00	2.00	[Ptp61F, sgg]
GO:0046626	regulation of insulin receptor signaling pathway	7.69	2.00	[Ptp61F, sNPF]
GO:0061319	nephrocyte differentiation	18.18	2.00	[numb, zfh1]
KEGG:00520	Amino sugar and nucleotide sugar metabolism	6.52	3.00	[Cht5, UGP, mmy]
Gene Network Results for Control & Diapause Fecundity				
GOID	GOTerm	% Associated Genes	Nr. Genes	Associated Genes Found
GO:0061343	cell adhesion involved in heart morphogenesis	15.38	2.00	[Gli, nrv2]
GO:0046112	nucleobase biosynthetic process	10.00	2.00	[Adk1, ade5]
GO:0004017	adenylate kinase activity	28.57	2.00	[Adk1, CG5626]
R-DME:166658	Complement cascade	9.52	2.00	[CG12374, CG18585]
R-DME:8852809	CPN, CPB2 cleave C3a, C5a	10.00	2.00	[CG12374, CG18585]
R-DME:977606	Regulation of Complement cascade	9.52	2.00	[CG12374, CG18585]
GO:0004181	metallocarboxypeptidase activity	8.70	2.00	[CG12374, CG18585]
GO:0045840	positive regulation of mitotic nuclear division	15.38	2.00	[APC7, Moe]
GO:0007368	determination of left/right symmetry	16.67	2.00	[Moe, puc]
GO:0042067	establishment of ommatidial planar polarity	5.48	4.00	[CadN2, Cct1, ara, puc]
GO:0016318	ommatidial rotation	6.25	2.00	[CadN2, puc]

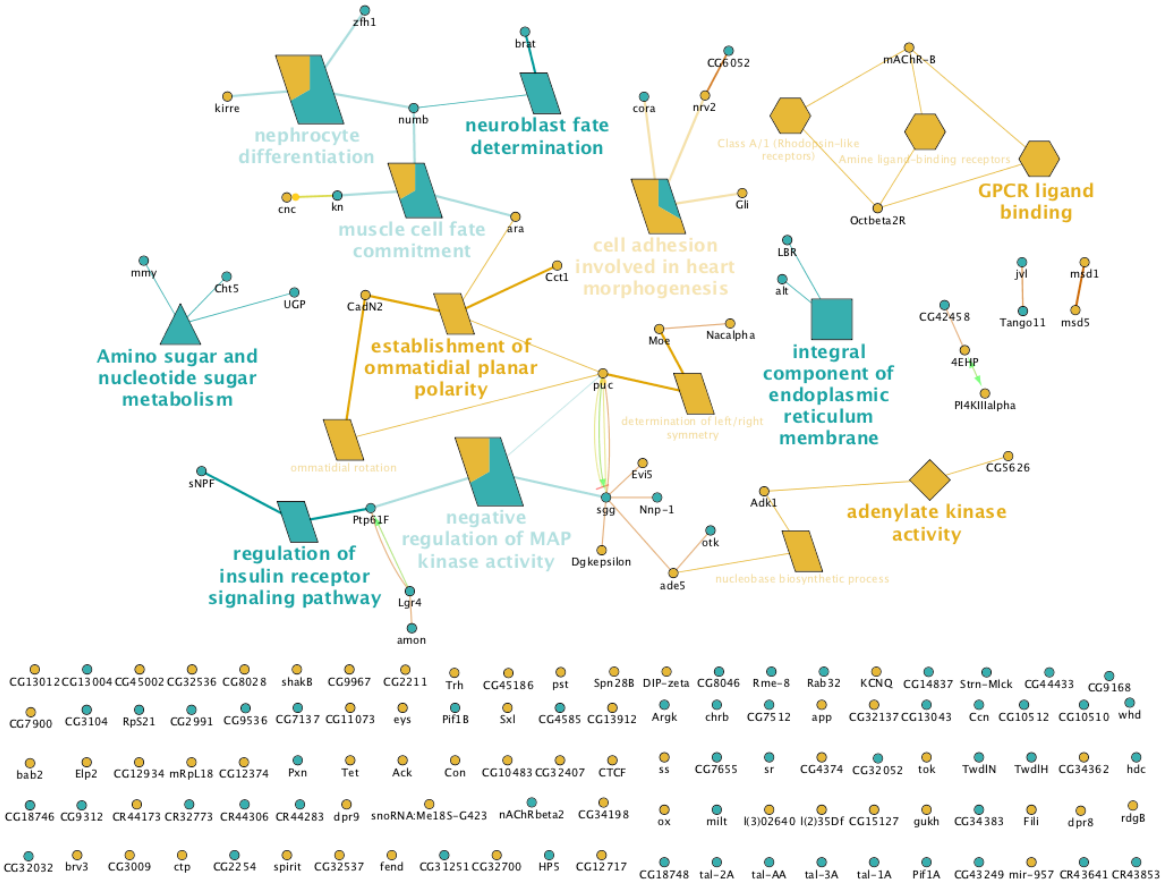

Figure 24 Gene Network of Combined Diapause GWAS Gene Candidate
Gene network overrepresentation analysis for all genes annotated in both diapause GWAS results. Circles are individual genes annotated in each the data set, color coded by GWAS analysis method (lifespan: blue, fecundity: yellow). Parallelograms show GO biological processes, squares show GO cellular components, diamonds depict GO molecular functions, hexagons indicate reactome pathways, and triangles depict KEGG pathways, represented in the gene set. Shapes are filled based on the contribution from either the post-diapause fecundity or total diapause lifespan gene set content. Green arrows indicate positive gene regulation, while orange lines indicate protein-protein interactions.

Table 5 Summary of Combined Diapause Gene Networks

GOID	GOTerm	% Associated Genes	Nr. Genes	Lifespan Genes	Fecundity Genes	%Genes Lifespan	%Genes Fecundity
GO:0061319	nephrocyte differentiation	27.27	3.00	[numb, zfh1]	[kirre]	66.67	33.33
GO:0061343	cell adhesion involved in heart morphogenesis	23.08	3.00	[cora]	[Gli, nrv2]	33.33	66.67
GO:0046626	regulation of insulin receptor signaling pathway	7.69	2.00	[Ptp61F, sNPF]	-	100.00	0.00
KEGG:00520	Amino sugar and nucleotide sugar metabolism	6.52	3.00	[Cht5, UGP, mmy]	-	100.00	0.00
GO:0007400	neuroblast fate determination	7.14	2.00	[brat, numb]	-	100.00	0.00
GO:0030176	integral component of endoplasmic reticulum membrane	6.25	2.00	[LBR, alt]	-	100.00	0.00
GO:0042693	muscle cell fate commitment	18.75	3.00	[kn, numb]	[ara]	66.67	33.33
GO:0043407	negative regulation of MAP kinase activity	30.00	3.00	[Ptp61F, sgg]	[puc]	66.67	33.33
GO:0046112	nucleobase biosynthetic process	10.00	2.00	-	[Adk1, ade5]	0.00	100.00
GO:0004017	adenylate kinase activity	28.57	2.00	-	[Adk1, CG5626]	0.00	100.00
R-DME:373076	Class A/1 (Rhodopsin-like receptors)	6.90	2.00	-	[Octbeta2R, mAChR-B]	0.00	100.00
R-DME:375280	Amine ligand-binding receptors	15.38	2.00	-	[Octbeta2R, mAChR-B]	0.00	100.00
R-DME:500792	GPCR ligand binding	6.25	2.00	-	[Octbeta2R, mAChR-B]	0.00	100.00
GO:0007368	determination of left/right symmetry	16.67	2.00	-	[Moe, puc]	0.00	100.00
GO:0042067	establishment of ommatidial planar polarity	5.48	4.00	-	[CadN2, Cct1, ara, puc]	0.00	100.00
GO:0016318	ommatidial rotation	6.25	2.00	-	[CadN2, puc]	0.00	100.00

To see which, if any, of the identified genes were also observed in other studies of diapause, an overlap analysis was enacted using data from two separate differential gene expression papers. One paper examined the difference in gene expression profiles between diapausing and non-diapausing flies using whole-body extracts (Kubrak et al. 2016), while the other examined two tissues separately, the ovary and the brain (Zhao et al. 2016). Seventeen genes were found in common between either of the two expression analyses and our study of diapause lifespan, while 22 were found in common for diapause fecundity. A subset of these overlapping genes was also represented in our gene network analysis and are presented in Table 6 on the following page.

Table 6 Subset of Genes from Overlap Analysis

Gene ID	Gene Name	Function
alt	aluminum tubes	Protein features are: Ribosome-binding protein 1.
ara	araucan	A homeodomain-containing protein of the TALE subfamily and one of the three components of the Iroquois gene complex. Acts as a transcriptional regulator to control territorial and cell fate specification, cell sorting behaviour and pattern formation. It also plays a transcription independent cell-autonomous role in cell proliferation control. [Date last reviewed: 2018-09-20]
cnc	cap-n-collar	Encodes a transcription factor that interacts with the product of Keap1 to regulate the activation of genes by oxidative stress. Also contributes to mRNA localization mediated by microtubules, dendrite morphogenesis and intestinal stem cell homeostasis. [Date last reviewed: 2018-03-08]
Gli	Gliotactin	Gliotactin is a transmembrane protein localized at tricellular junctions. It is necessary for septate junction and permeability barrier formation. [Date last reviewed: 2016-06-23]
jvl	Javelin-like	Javelin-like is a microtubule associated protein whose roles include actin and microtubule organization during oocyte development and bristle growth. [Date last reviewed: 2017-07-20]
kirre	Kin of irre	Kin of irre is a transmembrane protein of the Ig superfamily that is involved in heterotypic interactions with sns. Interactions with intracellular adaptors regulate cytoskeleton dynamics. Its biological roles include myoblast aggregation and fusion, nephrocyte diaphragm formation, and cell sorting in the eye and wing imaginal discs. [Date last reviewed: 2016-06-30]
Moe	Moesin	A Ezrin, Radixin, Moesin (ERM) protein involved in cortical cytoskeleton stability. Regulates crb and Rho1. Roles include apical-basal polarity, mitotic spindle organisation and epithelial integrity. [Date last reviewed: 2016-06-30]
nrv2	nervana 2	Sodium/potassium-transporting ATPase subunit beta
numb	numb	A membrane-associated inhibitor of Notch signaling. Controls neuroblast and sense organ precursor asymmetric division. It is asymmetrically localized during mitosis and segregates exclusively to one of two daughter cells. [Date last reviewed: 2017-02-09]
Octbeta2R	Octopamine β2 receptor	Protein features are: Basic-leucine zipper domain; G protein-coupled receptor, rhodopsin-like; GPCR, rhodopsin-like, 7TM
otk	Off-track	Associates with PlexA to receive a repulsive signal from Sema-1a contributing to axon guidance in the central nervous system and motor neurons. It is also associated with non-canonical Wnt signaling, opposing canonical Wnt signaling activation. [Date last reviewed: 2016-09-01]
Ptp61F	Protein-tyrosine phosphatase 61F	Protein-tyrosine phosphatase
sgg	shaggy	A Glycogen Synthase Kinase 3, and a key component of the β-catenin destruction complex. It functions in the canonical Wnt cascade. [Date last reviewed: 2016-06-30]

Discussion

Multiple studies of diapause across several taxa have shown that diapause is anything but simple quiescence, or state of inactivity. More evidence supports the idea that this phenotype is a complex trait actively initiated in response to unfavorable stimuli, and involves a substantial switch in organism metabolism, signaling and likely genetic architecture. Our study has revealed 151 candidate genes that may be involved in the diapause process and its regulation. Several genes returned from these analyses were also found by other groups and further, gene network analyses show enrichment in pathways/terms previously implicated by other studies of diapause across multiple species. Collectively, this indicates that our method is robust and that the identified genes have a high probability of being involved in the diapause phenotype

The representation of categories previously implicated in diapause, such as insulin receptor pathway regulation, negative MAPK regulation, and innate immunity, suggest that the GWAS was successful. The additional presence of the GPCR ligand binding category could implicate a potential hormone with the potential to regulate systemic changes that occur during diapause. Despite current efforts in this field, no one hormone is currently deemed responsible for diapause induction in *Drosophila melanogaster*. Furthermore, the gene candidates found to overlap with previous transcriptome analyses of diapause as well as those with multiple genetic variants associated, provide an excellent foundation of candidates to test functionally, narrowing down the pool from ~150 to about 20. One difficulty to overcome is testing those genes that lead to lethality when mutated, or affect the overall health of the fly regardless of environmental condition. It would be ideal to generate a system

in which gene function is only altered under diapause induction conditions (10˚C ; 8 light : 16 dark).

While diapause was protective for a majority of the DGRP lines, 14 of the lines benefited from diapause induction and significantly increased their lifespans post-diapause. This indicates that perhaps some of the genes involved in increasing longevity may have been permanently altered during diapause, and remain so afterward. It is interesting to note that many of the 14 DGRP lines that showed a significant increase in their PDLS, possessed shorter control lifespans, indicating that there may be a tradeoff between maximal lifespan under optimal growth conditions versus maximal lifespan under stressful conditions. Though many lines survived diapause conditions and possessed post-diapause lifespans comparable to their control lifespans, most of the strains were significantly less fecund post-diapause than their control counterparts. This brings to light the potential tradeoff between survival and reproductive ability, and raises the question of which is more beneficial, individual survival or the passing on of genetic information? It should be noted that diapausing conditions for this study were 5 weeks of induction, the typical length of winter, and the "protection" of the germline. To more thoroughly address this question, flies should be tested at 3 weeks of diapause after recovery, and compared to age matched controls, which may result in post-diapause fecundities more similar to controls. It has also been demonstrated that fertility decreases after removal from diapause, in proportion to their diapause duration (Tatar et al. 2001), which could be explained if any premature reinstatement of ovarian development removes the protection of the GSCs. The question remains as to whether the mechanism of this cell death and germ-line stem cell preservation, maintenance, and rejuvenation, is the

same as that observed in *C. elegans* (Fukuyama et al. 2012; Angelo and Van Gilst 2009; Seidel and Kimble 2011), as mentioned in the introduction.

While traditionally "quantified" using egg chamber development criteria, it has been shown in *D. melanogaster*, that delay in egg chamber development does not persist throughout the entire diapause period, but terminates following 4-6 weeks of diapause induction, with complete ovarian development seen by 8-13 weeks (Saunders et al. 1989; Kubrak et al. 2014). In addition, our studies indicate that during diapause the older vitellogenic egg chambers are degraded, and thus may still be present if degradation hasn't completely occurred (data not shown), leading to an incorrect classification of a fly as non-diapausing. As such, the experiments performed in this thesis utilized the difference between diapause lifespan and control lifespan, in addition to the difference between post-diapause fecundity and normal fecundity to quantify diapause induction. It would be interesting to see how our method of quantifying the diapause phenotype matches up to that typically used, by measuring proportion of pre-vitellogenic ovaries and then performing the GWAS analysis using the resultant data. Correlation analyses could then be used to see if the two methods of quantifying the diapause phenotype align, or if one is better suited for determination of diapause induction.

As with any experiment, these findings require replication and careful experimental control to confirm the validity of the results, and future work should aim to functionally analyze and confirm the involvement of the candidate genes found. Ideally a multi-phenotype GWAS should be performed to further solidify, and narrow down potential candidates. Platforms exist that allow users to input multiple quantitative phenotypes for GWAS, which would enable the marriage of the two sub-phenotypes affected by diapause, potentially

increasing the power of the GWAS of diapause. Studies in *Drosophila* offer great potential to probe unanswered questions regarding aging and its associated degeneration of the body using diapause as a model. Given that the proportion of the world's population will almost double from 12 to 22 percent over the next 35 years (World Health Organization 2018), it is my strong opinion, that current research should not seek to further extend the human lifespan, but instead improve the quality of our extended "golden years". By ameliorating and delaying the degenerative processes associated with aging, we can give back the independence, health, and happiness, to those suffering from age related diseases. Aging, it is a matter of life and death.

Supplemental Materials

All Analysis Methods with Top Associated Genes -

https://drive.google.com/open?id=1XC5xmHbehDsPzKsTt_qiAjqVnYIeNxeq

Statistical Analysis Results & Individual Line Statistics

https://drive.google.com/open?id=1GBOeoh98znhvJ30ocBqYJaDQj53QNZIL

Bibliography

Allen, M.J. 2007. What makes a fly enter diapause? *Fly* 1(6), pp. 307–310.

Anduaga, A.M., Nagy, D., Costa, R. and Kyriacou, C.P. 2018. Diapause in Drosophila melanogaster - Photoperiodicity, cold tolerance and metabolites. *Journal of Insect Physiology* 105, pp. 46–53.

Angelo, G. and Van Gilst, M.R. 2009. Starvation protects germline stem cells and extends reproductive longevity in C. elegans. *Science* 326(5955), pp. 954–958.

Anon 2015. *World Population Ageing 2015*. United Nations.

Baker, D.A. and Russell, S. 2009. Gene expression during Drosophila melanogaster egg development before and after reproductive diapause. *BMC Genomics* 10, p. 242.

Barja, G. 2002. Rate of generation of oxidative stress-related damage and animal longevity. *Free Radical Biology & Medicine* 33(9), pp. 1167–1172.

Barzilai, N., Huffman, D.M., Muzumdar, R.H. and Bartke, A. 2012. The critical role of metabolic pathways in aging. *Diabetes* 61(6), pp. 1315–1322.

Baugh, L.R. 2013. To grow or not to grow: nutritional control of development during Caenorhabditis elegans L1 arrest. *Genetics* 194(3), pp. 539–555.

Braeckman, B.P., Houthoofd, K., De Vreese, A. and Vanfleteren, J.R. 1999. Apparent uncoupling of energy production and consumption in long-lived Clk mutants of Caenorhabditis elegans. *Current Biology* 9(9), pp. 493–496.

Broughton, S. and Partridge, L. 2009. Insulin/IGF-like signalling, the central nervous system and aging. *The Biochemical Journal* 418(1), pp. 1–12.

Budi, E.H., Muthusamy, B.-P. and Derynck, R. 2015. The insulin response integrates increased TGF-β signaling through Akt-induced enhancement of cell surface delivery of TGF-β receptors. *Science Signaling* 8(396), p. ra96.

Bush, W.S. and Moore, J.H. 2012. Chapter 11: Genome-wide association studies. *PLoS Computational Biology* 8(12), p. e1002822.

Butcher, R.A., Ragains, J.R., Li, W., Ruvkun, G., Clardy, J. and Mak, H.Y. 2009. Biosynthesis of the Caenorhabditis elegans dauer pheromone. *Proceedings of the National Academy of Sciences of the United States of America* 106(6), pp. 1875–1879.

Copeland, J.M., Cho, J., Lo, T., et al. 2009. Extension of Drosophila life span by RNAi of the mitochondrial respiratory chain. *Current Biology* 19(19), pp. 1591–1598.

Davy, P.M.C., Allsopp, R.C., Donlon, T.A., Morris, B.J., Willcox, D.C. and Willcox, B.J. 2018. FOXO3 and exceptional longevity: insights from hydra to humans. *Current Topics in Developmental Biology* 127, pp. 193–212.

De Meyts, P. 2000. The Insulin Receptor and Its Signal Transduction Network. In: De Groot,

L. J., Chrousos, G., Dungan, K., et al. eds. *Endotext*. South Dartmouth (MA): MDText.com, Inc.

Denlinger, D.L. 2002. Regulation of diapause. *Annual Review of Entomology* 47, pp. 93–122.

Ellis, L.L., Huang, W., Quinn, A.M., et al. 2014. Intrapopulation genome size variation in D. melanogaster reflects life history variation and plasticity. *PLoS Genetics* 10(7), p. e1004522.

Epel, E.S. and Lithgow, G.J. 2014. Stress biology and aging mechanisms: toward understanding the deep connection between adaptation to stress and longevity. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 69 Suppl 1, pp. S10-6.

Fedarko, N.S. 2018. Theories and mechanisms of aging. In: Reves, J. G., Barnett, S. R., McSwain, J. R., and Rooke, G. A. eds. *Geriatric Anesthesiology*. Cham: Springer International Publishing, pp. 19–25.

Fielenbach, N. and Antebi, A. 2008. C. elegans dauer formation and the molecular basis of plasticity. *Genes & Development* 22(16), pp. 2149–2165.

Fontana, L., Partridge, L. and Longo, V.D. 2010. Extending healthy life span--from yeast to humans. *Science* 328(5976), pp. 321–326.

Fukuyama, M., Sakuma, K., Park, R., et al. 2012. C. elegans AMPKs promote survival and arrest germline development during nutrient stress. *Biology open* 1(10), pp. 929–936.

Goldsmith, T.C. 2014. Modern evolutionary mechanics theories and resolving the programmed/non-programmed aging controversy. *Biochemistry. Biokhimiia* 79(10), pp. 1049–1055.

Guo, X. and Wang, X.-F. 2009. Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Research* 19(1), pp. 71–88.

Hand, S.C., Denlinger, D.L., Podrabsky, J.E. and Roy, R. 2016. Mechanisms of animal diapause: recent developments from nematodes, crustaceans, insects, and fish. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 310(11), pp. R1193-211.

Houtkooper, R.H., Williams, R.W. and Auwerx, J. 2010. Metabolic networks of longevity. *Cell* 142(1), pp. 9–14.

Huang, W., Massouras, A., Inoue, Y., et al. 2014. Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. *Genome Research* 24(7), pp. 1193–1208.

Hwang, A.B., Jeong, D.-E. and Lee, S.-J. 2012. Mitochondria and organismal longevity. *Current Genomics* 13(7), pp. 519–532.

Hwangbo, D.S., Gershman, B., Tu, M.-P., Palmer, M. and Tatar, M. 2004. Drosophila dFOXO controls lifespan and regulates insulin signalling in brain and fat body. *Nature* 429(6991), pp. 562–566.

Ivanov, D.K., Escott-Price, V., Ziehm, M., et al. 2015. Longevity GWAS using the

drosophila genetic reference panel. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 70(12), pp. 1470–1478.

Jeong, P.-Y., Jung, M., Yim, Y.-H., et al. 2005. Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone. *Nature* 433(7025), pp. 541–545.

Jin, K. 2010. Modern Biological Theories of Aging. *Aging and disease* 1(2), pp. 72–74.

Johnson, S.C., Rabinovitch, P.S. and Kaeberlein, M. 2013. mTOR is a key modulator of ageing and age-related disease. *Nature* 493(7432), pp. 338–345.

Kang, D.S., Cotten, M.A., Denlinger, D.L. and Sim, C. 2016. Comparative Transcriptomics Reveals Key Gene Expression Differences between Diapausing and Non-Diapausing Adults of Culex pipiens. *Plos One* 11(4), p. e0154892.

Kankare, M., Parker, D.J., Merisalo, M., Salminen, T.S. and Hoikkala, A. 2016. Transcriptional Differences between Diapausing and Non-Diapausing D. montana Females Reared under the Same Photoperiod and Temperature. *Plos One* 11(8), p. e0161852.

Kankare, M., Salminen, T., Laiho, A., Vesala, L. and Hoikkala, A. 2010. Changes in gene expression linked with adult reproductive diapause in a northern malt fly species: a candidate gene microarray study. *BMC Ecology* 10, p. 3.

Karp, X. 2018. Working with dauer larvae. *Wormbook: the Online Review of C. Elegans Biology* 2018, pp. 1–19.

Kenyon, C. 2011. The first long-lived mutants: discovery of the insulin/IGF-1 pathway for ageing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366(1561), pp. 9–16.

Kenyon, C.J. 2010. The genetics of ageing. *Nature* 464(7288), pp. 504–512.

Kim, K., Sato, K., Shibuya, M., et al. 2009. Two chemoreceptors mediate developmental effects of dauer pheromone in C. elegans. *Science* 326(5955), pp. 994–998.

Kimura, K.D., Tissenbaum, H.A., Liu, Y. and Ruvkun, G. 1997. daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science* 277(5328), pp. 942–946.

Klass, M.R. 1983. A method for the isolation of longevity mutants in the nematode Caenorhabditis elegans and initial results. *Mechanisms of Ageing and Development* 22(3–4), pp. 279–286.

Klein, C., Lohmann, K. and Ziegler, A. 2012. The Promise and Limitations of Genome-wide Association Studies. *The Journal of the American Medical Association* 308(18), p. 1867.

Kostál, V. 2006. Eco-physiological phases of insect diapause. *Journal of Insect Physiology* 52(2), pp. 113–127.

Kubrak, O.I., Kučerová, L., Theopold, U. and Nässel, D.R. 2014. The sleeping beauty: how reproductive diapause affects hormone signaling, metabolism, immune response and somatic maintenance in Drosophila melanogaster. *Plos One* 9(11), p. e113051.

Kubrak, O.I., Kučerová, L., Theopold, U., Nylin, S. and Nässel, D.R. 2016. Characterization of Reproductive Dormancy in Male Drosophila melanogaster. *Frontiers in physiology* 7, p. 572.

Kučerová, L., Kubrak, O.I., Bengtsson, J.M., et al. 2016. Slowed aging during reproductive dormancy is reflected in genome-wide transcriptome changes in Drosophila melanogaster. *BMC Genomics* 17, p. 50.

Laplante, M. and Sabatini, D.M. 2009. mTOR signaling at a glance. *Journal of Cell Science* 122(Pt 20), pp. 3589–3594.

Lee, R. and Mason, A. 2017. Cost of Aging. *Finance & development* 54(1), pp. 7–9.

Lee, S.S., Lee, R.Y.N., Fraser, A.G., Kamath, R.S., Ahringer, J. and Ruvkun, G. 2003. A systematic RNAi screen identifies a critical role for mitochondria in C. elegans longevity. *Nature Genetics* 33(1), pp. 40–48.

Li, X., Fang, P., Mai, J., Choi, E.T., Wang, H. and Yang, X. 2013. Targeting mitochondrial reactive oxygen species as novel therapy for inflammatory diseases and cancers. *Journal of hematology & oncology* 6, p. 19.

Liao, S., Broughton, S. and Nässel, D.R. 2017. Behavioral Senescence and Aging-Related Changes in Motor Neurons and Brain Neuromodulator Levels Are Ameliorated by Lifespan-Extending Reproductive Dormancy in Drosophila. *Frontiers in Cellular Neuroscience* 11, p. 111.

Liu, F., Li, K., Cai, W., Zhao, J., Zou, Y. and Hua, H. 2017. Knockdown of TOR causing ovarian diapause in a genetically stable brachypterous strain of Nilaparvata lugens. *Archives of Insect Biochemistry and Physiology*.

Liu, T., Zimmerman, K.K. and Patterson, G.I. 2004. Regulation of signaling genes by TGFbeta during entry into dauer diapause in C. elegans. *BMC Developmental Biology* 4, p. 11.

Liu, X., Jiang, N., Hughes, B., Bigras, E., Shoubridge, E. and Hekimi, S. 2005. Evolutionary conservation of the clk-1-dependent mechanism of longevity: loss of mclk1 increases cellular fitness and lifespan in mice. *Genes & Development* 19(20), pp. 2424–2434.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M. and Kroemer, G. 2013. The hallmarks of aging. *Cell* 153(6), pp. 1194–1217.

López-Otín, C., Galluzzi, L., Freije, J.M.P., Madeo, F. and Kroemer, G. 2016. Metabolic control of longevity. *Cell* 166(4), pp. 802–821.

Lumme, J., Oikarinen, A., Lakovaara, S. and Alatalo, R. 1974. The environmental regulation of adult diapause in Drosophila littoralis. *Journal of Insect Physiology* 20(10), pp. 2023–2033.

Mackay, T.F.C. and Huang, W. 2018. Charting the genotype-phenotype map: lessons from the Drosophila melanogaster Genetic Reference Panel. *Wiley interdisciplinary reviews. Developmental biology* 7(1).

Mackay, T.F.C., Richards, S., Stone, E.A., et al. 2012. The Drosophila melanogaster Genetic Reference Panel. *Nature* 482(7384), pp. 173–178.

MacRae, T.H. 2010. Gene expression, metabolic regulation and stress tolerance during diapause. *Cellular and Molecular Life Sciences* 67(14), pp. 2405–2424.

Martins, R., Lithgow, G.J. and Link, W. 2016. Long live FOXO: unraveling the role of FOXO proteins in aging and longevity. *Aging Cell* 15(2), pp. 196–207.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., et al. 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews. Genetics* 9(5), pp. 356–369.

Mookerjee, S.A., Divakaruni, A.S., Jastroch, M. and Brand, M.D. 2010. Mitochondrial uncoupling and lifespan. *Mechanisms of Ageing and Development* 131(7–8), pp. 463–472.

Newgard, C.B. and Pessin, J.E. 2014. Recent progress in metabolic signaling pathways regulating aging and life span. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 69 Suppl 1, pp. S21-7.

Pal, S. and Tyler, J.K. 2016. Epigenetics and aging. *Science advances* 2(7), p. e1600584.

Ragland, G.J., Denlinger, D.L. and Hahn, D.A. 2010. Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. *Proceedings of the National Academy of Sciences of the United States of America* 107(33), pp. 14909–14914.

Roser, M. and Ritchie, H. 2018. Burden of Disease - Our World in Data [Online]. Available at: https://ourworldindata.org/burden-of-disease [Accessed: 31 August 2018].

Salminen, T.S., Vesala, L., Laiho, A., Merisalo, M., Hoikkala, A. and Kankare, M. 2015. Seasonal gene expression kinetics between diapause phases in Drosophila virilis group species and overwintering differences between diapausing and non-diapausing females. *Scientific reports* 5, p. 11197.

Saunders, D.S., Henrich, V.C. and Gilbert, L.I. 1989. Induction of diapause in Drosophila melanogaster: photoperiodic regulation and the impact of arrhythmic clock mutations on time measurement. *Proceedings of the National Academy of Sciences of the United States of America* 86(10), pp. 3748–3752.

Saunders, D.S., Richard, D.S., Applebaum, S.W., Ma, M. and Gilbert, L.I. 1990. Photoperiodic diapause in Drosophila melanogaster involves a block to the juvenile hormone regulation of ovarian maturation. *General and Comparative Endocrinology* 79(2), pp. 174–184.

Schiesari, L., Andreatta, G., Kyriacou, C.P., O'Connor, M.B. and Costa, R. 2016. The Insulin-Like Proteins dILPs-2/5 Determine Diapause Inducibility in Drosophila. *Plos One* 11(9), p. e0163680.

Schiesari, L., Kyriacou, C.P. and Costa, R. 2011. The hormonal and circadian basis for insect photoperiodic timing. *FEBS Letters* 585(10), pp. 1450–1460.

Schiesari, L. and O'Connor, M.B. 2013. Diapause: delaying the developmental clock in

response to a changing environment. *Current Topics in Developmental Biology* 105, pp. 213–246.

Seidel, H.S. and Kimble, J. 2011. The oogenic germline starvation response in C. elegans. *Plos One* 6(12), p. e28074.

Sergiev, P.V., Dontsova, O.A. and Berezkin, G.V. 2015. Theories of aging: an ever-evolving field. *Acta naturae* 7(1), pp. 9–18.

Sim, C. and Denlinger, D.L. 2008. Insulin signaling and FOXO regulate the overwintering diapause of the mosquito Culex pipiens. *Proceedings of the National Academy of Sciences of the United States of America* 105(18), pp. 6777–6781.

Sim, C. and Denlinger, D.L. 2013. Insulin signaling and the regulation of insect diapause. *Frontiers in physiology* 4, p. 189.

Sommer, R.J. and Ogawa, A. 2011. Hormone signaling and phenotypic plasticity in nematode development and evolution. *Current Biology* 21(18), pp. R758-66.

Spielman, A. 1974. Effect of synthetic juvenile hormone on ovarian diapause of culex pipiens mosquitoes1. *Journal of Medical Entomology* 11(2), pp. 223–225.

Stranger, B.E., Stahl, E.A. and Raj, T. 2011. Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 187(2), pp. 367–383.

Tammariello, S.P. 2001. Regulation of the cell cycle during diapause. In: *Insect timing: circadian rhythmicity to seasonality*. Elsevier, pp. 173–183.

Tatar, M. 2010. Reproductive aging in invertebrate genetic models. *Annals of the New York Academy of Sciences* 1204, pp. 149–155.

Tatar, M., Chien, S.A. and Priest, N.K. 2001. Negligible Senescence during Reproductive Dormancy in Drosophila melanogaster. *The American Naturalist* 158(3), pp. 248–258.

Tatar, M. and Yin, C. 2001. Slow aging during insect reproductive diapause: why butterflies, grasshoppers and flies are like worms. *Experimental Gerontology* 36(4–6), pp. 723–738.

World Health Organization 2018. *World health statistics 2018: monitoring health for the SDGs, sustainable development goals*. Geneva: World Health Organization.

Xu, W.-H., Lu, Y.-X. and Denlinger, D.L. 2012. Cross-talk between the fat body and brain regulates insect developmental arrest. *Proceedings of the National Academy of Sciences of the United States of America* 109(36), pp. 14687–14692.

Xue, G., Restuccia, D.F., Lan, Q., et al. 2012. Akt/PKB-mediated phosphorylation of Twist1 promotes tumor metastasis via mediating cross-talk between PI3K/Akt and TGF-β signaling axes. *Cancer discovery* 2(3), pp. 248–259.

Yamamoto, R. and Tatar, M. 2011. Insulin receptor substrate chico acts with the transcription factor FOXO to extend Drosophila lifespan. *Aging Cell* 10(4), pp. 729–732.

Zhao, X., Bergland, A.O., Behrman, E.L., Gregory, B.D., Petrov, D.A. and Schmidt, P.S. 2016. Global Transcriptional Profiling of Diapause and Climatic Adaptation in Drosophila

melanogaster. *Molecular Biology and Evolution* 33(3), pp. 707–720.

Zhou, K.I., Pincus, Z. and Slack, F.J. 2011. Longevity and stress in Caenorhabditis elegans. *Aging* 3(8), pp. 733–753.

Zonato, V., Collins, L., Pegoraro, M., Tauber, E. and Kyriacou, C.P. 2017. Is diapause an ancient adaptation in Drosophila? *Journal of Insect Physiology* 98, pp. 267–274.